THE STRUCTURE AND PARAGRAPHING OF EXPOSITORY PROSE : A TRADITIONAL-RHETORICAL APPROACH WITH REFERENCE TO THE TEACHING OF WRITING

Gillian Daphne Elizabeth
Schärer

Ph.D.
University of Edinburgh
1980

ABSTRACT

This thesis originated in the writing problems of Malawian students learning through the medium of English. In most cases these students can perform adequately at sentence level but are unable to produce satisfactory expository text, a situation which often has serious educational consequences.

Traditional rhetoric is potentially a valuable tool for the teacher grappling with this problem, but its piecemeal formulation and unsystematic approach has led to its current disfavour in the British classroom. It needs, therefore, to be systematised and placed within a more unified theory, and for this purpose the experimental method is necessary to validate or otherwise certain commonly held prescriptive beliefs which seem to be intuitively correct. An attempt is made here to do just this with respect to the paragraph, this inevitably also involving a consideration of expository text structure in general.

The work begins with a review of the traditional rhetorical approach to the paragraph during the last hundred years. It goes on to consider the work of certain American teachers of college composition, in particular Frances Christensen and Alton Becker, who are basically traditional in their approach but

have at the same time been influenced by the trends in linguistics evident in the sixties.

The second and main part is experimental. An expository text is viewed as consisting of 'information blocks' of various sizes and at different levels, with the paragraph as a surface phenomenon imposed on this structure. Thus, while the main object of study is the orthographic 'paragraph', it must be approached via the more basic 'information block'. Experiments are carried out concerning sentence length (the sentence being viewed as an information block in its own right), the signalling function of the short sentence, the relationship between the divisions of a text into information blocks and its hierarchical structure, and the influence of format on paragraph divisions. In addition, the topic sentence, the bridge sentence, and the effectiveness of various types of signal to indicate transitions between information blocks are also touched on.

The final two chapters summarise the theoretical conclusions and attempt to relate the findings to the teaching of writing skills at tertiary level.

ACKNOWLEDGMENTS

I should first like to thank the Ministry of Overseas Development (as it was named in 1976) and the Free University of West Berlin, without whose generous financial support this thesis could not have been written.

Appreciation should also be expressed for the cooperation of those lecturers at Moray House College of Education, Edinburgh, and Chancellor College, University of Malawi who acted as patient and uncomplaining subjects, not to mention the many friends and relations involved in the pilot studies. Special mention must also be made of Mike Caslake and Guy Walker of the Chancellor College Geography Department, whose helpful supply of student essays and project work acted as an important initial stimulus.

Above all, however, I should like to thank my supervisor Professor S. Pit Corder for his never-failing support and encouragement, a major factor in the completion of this work.

DECLARATION

This thesis is my own work and composition

28/2/1980

TABLE OF CONTENTS

CHAPTER 1 : Preliminary Considerations : Problems and Aims

1 Introduction 1

2 English in Malawi 6

2.1 The role of English as a Second Language 6

2.2 English within the school system 7

3 Wider influences 13

3.1 Linguistics & psychology 13

3.2 Literature & creative writing 16

4 A possible answer : ESP 17

5 General aims 19

6 Some definitions 21

6.1 Expository prose 22

6.2 Discourse, text & passage 24

6.3 Paragraph and information block 25

7 Overall plan & hypotheses 25

7.1 Chapters 2 to 3 25

7.2 Chapters 4 to 8 26

7.3 Chapters 9 and 10 27

CHAPTER 2 : The Legacy of Traditional Rhetoric

1 Introduction 28

2 Historical background 29

2.1 Aristotelian rhetoric 29

2.2 The last hundred years 30

2.3 A theory of the paragraph 33

3 Survey of representative texts 34

3.1 Definitions of the paragraph 35

3.2 Bain's Six Rules 40

3.3 Genung's Three Principles 41

3.4 The topic sentence 48

3.5 Paragraph development 51

3.6 Sentence length 54

3.7 Paragraph length 57

4 Evaluation 59

CHAPTER 3 : The Contribution of American Teachers of College Composition in the Sixties

1 Introduction 61

2 Francis Christensen 62

2.1 Christensen versus Hunt : the cumulative sentence 63

2.2 Application of the cumulative sentence to the paragraph 66

2.3 Evaluation 72

3 David Karrfalt 77

3.1 Revision of Christensen's model 77

3.2 Evaluation 79

4 Willis Pitkins 80

4.1 Pitkins' model 81

4.2 Evaluation 84

5 Robert Kaplan 86

5.1 Kaplan's model 86

5.2 Evaluation 93

6 A.E. Jones & C.W.Faulkner 94

6.1 Jones & Faulkner's model 95

6.2 Evaluation 96

7 Alton Becker 106

7.1 Becker's theory of paragraph analysis 106

7.2 Evaluation 107

7.3 Field analysis of the paragraph 115

7.4 Becker's main contributions to the theory of the paragraph 117

8 Conclusion 123

CHAPTER 4 : Towards a Theory of the Paragraph

1 Later developments and their limitations 127

2 The paragraph in traditional rhetoric 128

3 A possible model of expository prose structure 132

4 Experimental evidence 134

CHAPTER 5 : The Significance of Sentence Length

1 Introduction 138

2 Sentence length and the second-language learner 139

2.1 Optimum sentence length : the psycholinguistic approach 140

2.2 Optimum sentence length : the language-teacher's approach 142

2.3 Definitions of 'long' and 'short' 144

3 The experiments 145

3.1 Method 146

3.2 Subjects 146

3.3 Material 147

3.4 Results : statistical analyses 156

4 General discussion of results 174

4.1 Experiment 1 : Hypothesis 1a 174

4.2 Hypothesis 1b 177

4.3 Experiment 2 : Hypothesis 2 178

CHAPTER 6 : The Information Block

1 Introduction 179

1.1 The information block and the paragraph 179

1.2 The investigation 180

2 The analysis according to signals identified in the text 181

2.1 Extrinsic signalling 184

2.2 Intrinsic signalling 194

3 Grimes' Semantic Grammar of Propositions as a model for text chunking 200

3.1 Description 201

3.2 Application 205

3.3 The Vertical Movement Index 207

4 Experiment 3 208

4.1 Subjects and method 210

4.2 Raw results 211

4.3 Statistical analysis 220

4.4 General discussion of results 228

5 A comparison with Koen, Becker and Young's experiment 238

CHAPTER 7 : The Bridge Sentence and the Topic Sentence

1 Introduction 241

2 The bridge sentence 241

2.1 Bridge sentences in the experimental passages 242

2.2 Rees & Urquhart's examples 243

2.3 Application to the experimental passages 252

2.4 Bridge sentences in general 259

3 The topic sentence 260

3.1 Some possible definitions 261

3.2 An alternative definition 262

3.3 Topic sentences as macro-propositions 264

3.4 The length of the topic sentence 268

CHAPTER 8 : Paragraph Length

1 Introduction 272

2 D.S. Knapp's experiment 273

3 Experiment 4 275

3.1 Subjects 276

3.2 Material 276

3.3 Method 278

3.4 Statistical analysis of results 279

3.5 Interpretation of results 288

CHAPTER 9 : Theoretical Summary

1 Introduction 294

2 Some difficulties resulting from the use of text 294

3 The experimental work 297

3.1 Experiments 1 and 2 297

3.2 Experiment 3 299

3.3 Experiment 4 302

4 The bridge sentence and the topic sentence 303

4.1 The bridge sentence 304

4.2 The topic sentence 304

5 An analysis of two representative samples of student writing 306

5.1 Passage 1 310

5.2 Passage 2 312

5.3 The approach of these writers to text chunking 316

CHAPTER 10 : Some Pedagogical Applications

1 Introduction 318

2 Sentence length : Experiments 1 and 2 318

3 The paragraph in relation to the hierarchical structure of the text : Experiments 3 and 4 321

4 Extrinsic linking devices : Experiment 3 326

5 Some more general implications 328

6 An area for future research : process rather than product 331

BIBLIOGRAPHY 333

APPENDICES :

APPENDIX 1	Fifteen representative textbooks of traditional rhetoric, 1866-1978.	344
APPENDIX 2a	Passages used for Experiment 1.	347
APPENDIX 2b	Passages used for Experiment 2.	354
APPENDIX 2c	Experiments 1 and 2 : modified kernels.	348
APPENDIX 2d	Experiment 2 : content structure of the three passages.	361
APPENDIX 3	Experiment 1 : statistical analyses	364
APPENDIX 4	Experiment 3 : Passage A, Passage B and Passage C.	369
APPENDIX 5	Experiment 3 : Lexical cohesion chains and movements from particular to general in the three passages.	386
APPENDIX 6	Experiment 3 : An analysis of the three passages according to a modified version of Grimes' Semantic Grammar of Propositions.	392
APPENDIX 7	Experiment 3 : instructions and raw results.	396
APPENDIX 8	Experiment 3 : statistical analyses.	403
APPENDIX 9	Textbooks used for the analysis of theme sentence lengths.	409
APPENDIX 10	Experiment 4 : the experimental passages	411
APPENDIX 11	Experiment 4 : statistical analyses	426

FIGURES

1	Example of student writing : fourth year secondary	3
2	Example of student writing : second year tertiary	5
3	Britton's three functional types	22
4	Christensen's analysis (1)	69
5	Christensen's analysis (2)	71
6	Christensen's analysis : the Soil Erosion passage	74
7	Karrfalt's analysis	78
8	Pitkins' analysis	83
9	Pitkin's analysis : the Soil Erosion passage	85
10	Kaplan's analysis	88
11	Kaplan's analysis : the Soil Erorion passage	89
12	Simplified tree diagram of Kaplan's analysis (the Soil Erorion passage)	92
13	Jones and Faulkner's five types of paragraph	98
14	Jones and Faulkner's analysis : the Soil Erosion passage	102
15	Hierarchical tree of Pitkin's model : the Soil Erosion passage	103
16	Hierarchical tree of Jones and Faulkner's model : the Soil Erosion passage	104
17	Becker's analysis (1) The TRI pattern	108
18	Becker's analysis (2) The QA pattern	109
19	Becker's analysis : the Soil Erosion passage	112

20	Short sentences in the three experimental passages	188
21	Experiment 3 : results	212
22	Experiment 3 : weighted choices	218
23	Rees and Urquhart's passages	244
24	Analysis of Rees & Urquhart's bridge sentences	251
25	Passage A Sentence 8 : analysis	253
26	Passage B Sentence 33 : analysis	256
27	Passage B Sentence 24 : analysis	258
28	Topic sentences, Passages A, B and C	265
29	Experiment 4 : Adjusted residuals : Formats plotted against versions	289
30	Experiment 4 : Adjusted residuals : versions plotted against formats	290
31	Tree diagram of an example of student writing (Passage 1)	308
32	Tree diagram of an example of student writing (Passage 2)	309

TABLES

1	Definitions of the paragraph 1866-1978	35
2	Definitions of the paragraph 1866-1978 : common elements	38
3	Bain's Six Rules and Genung's Three Principles 1866-1978	44
4	Jones & Faulkner's meaning relationships	97
5	Jones & Faulkner's higher level classes	97
6	Becker's model : field analysis of the paragraph	116
7	Lexical equivalence chains : the English Constitution passage	119
8	Comparison between sentences written by one second language and seven mother-tongue speakers	139
9	Passages for Experiment 1 : total words	151
10	Passages for Experiment 1 : total sentences	151
11	Passages for Experiment 1 : average sentence lengths	151
12	Experiment 1 : raw results	156
13	Experiment 2 : raw results	157
14	Experiment 1 : the Chi-Square Test applied to each passage	159
15	Experiment 1 : the Chi-Square Test applied to each sentence length	163
16	Experiment 1 : adjusted residuals	166
17	Experiment 1 : the Kolmogorov-Smirnov One-Sample Test	168

18	Experiment 1 : Somer's 'd'	170
19	Experiment 1 : The relative positions of varied and medium length sentences	171
20	Experiment 2 : Chi-Square Test : the positioning of short sentences	173
21	The readability of the passages, as measured by two readability formulae	176
22	Position of signals and boundary choices	185
23	Vertical Movement Index measurements for each passage	208
24	Numbers of markers (all types) found at different divisions	224
25	Numbers of intrinsic, extrinsic and VMI markers found at different divisions	226
26	Relationships between specific markers and subjects' choices	227
27	Pattern of distribution of VMI measurements over all three passages	229
28	Pattern of distribution of number of propositions at each division not accompanied by an upward movement	232
29	Distribution of intrinsic and extrinsic signals over boundaries chosen by high percentages of subjects	236
30	Possible bridge sentences in the experimental passages	243
31	Bridge sentences : results of Rees and Urquhart's experiment	246
32	Analysis of Rees and Urquhart's bridge sentences	252
33	Analysis of theme sentence lengths in ten selected text books	269

34	The functions of the short sentence in the experimental passages	271
35	The results of Knapp's experiment on paragraph length	274
36	Experiment 4 : raw results	279
37	Experiment 4 : Chi-Square Tests applied to each format	280
38	Experiment 4 : Formats : Adjusted residuals	282
39	Experiment 4 : Chi-Square Tests applied to each version	284
40	Experiment 4 : Versions : Adjusted residuals	286
41	Experiment 4 : Goodman and Kruskal's Lamda measure of association	288

CHAPTER 1

PRELIMINARY CONSIDERATIONS : PROBLEMS AND AIMS

1. Introduction

The problem that has led to the investigation pursued in these pages is one familiar to all teachers of tertiary level students learning through the medium of English, but for whom English is not the mother tongue. It may indeed be a problem occasionally faced by teachers of British students in sixth forms and institutes of higher education. This is the inability of some students to produce coherent and well-organised expository writing, even when they have a reasonable command of the language at sentence level.

Such a performance deficit first came to the serious attention of the writer between 1972 and 1976, while teaching courses in English for Academic Purposes at the University of Malawi (Chancellor College). The standard of English in Malawian secondary schools was high compared to other countries where similar conditions prevailed. Nevertheless it seemed that many students, even after twelve years of school English, which included six years in which English had been used as the medium of instruction, were unable to cope with the more rigorous demands of the university. A particular cause for concern was that there was no way of telling how many students did not perform to the level of their real ability due to inadequacy in written expression.

The situation has so far been presented in vague terms only, and some examples are necessary. Two typical pieces of student writing can be seen in Figures 1 and 2. Neither of these examples was written for the English teacher, so neither would have been judged on the competence of its English expression. However, although no mark would have been given for the language of the essay, the subject teacher would undoubtedly have been influenced, consciously or unconsciously, by the language chosen to convey the writer's ideas. Figure 1 is not a university essay, being the written part of an independent project produced for the Malawi School Certificate Geography Examination. The standard is about average or slightly below average, so the writer may or may not have gone on to some form of higher education. This piece of work, however, is not noticeably different from that of the poorest Diploma in Education students when they begin their studies. Figure 2 is an essay written under examination conditions by a Diploma in Education student after two years of higher study (including regular English classes). The subject is the Sociology of Education. This work can, again, be judged as average to below average for the group as a whole.

These essays will be discussed in more detail later, but a superficial reading will be enough to show that something is wrong with each, above and beyond the usual syntactic and lexical errors. This weakness lies in the area of organisation and coherence, the area once classified under the heading of "rhetoric". There are inadequacies in overall planning, paragraph organisation and development, the distribution of

long and short sentences, cohesion and coherence, and the distribution of given and new information within individual sentences, to mention only the main categories. It is significant that none of these areas are usually represented in error analyses, except for specific errors that can easily be pinpointed, such as faulty reference or an incorrectly used logical connector. An exception is Bhatia (1974), who divides errors into those concerned with mechanics (ie. errors of a grammatical nature within the sentence) and those concerned with organisation (errors above sentence level). However, she makes no attempt to ennumerate organisational errors, simply awarding a subjective "yes" or "no" to each of her organisational sub-categories.

FIGURE 1 EXAMPLE OF STUDENT WRITING : FOURTH YEAR SECONDARY

Produce on Sale at Malembo Local Market

1 [1]Malembo is a local market in the Central Region, Lilongwe District, and it is situated in the western side of Lilongwe town. [2]It is in a rural area thirty-five miles away from the town. (See Map 1).

2 [3]The market is near the main road so that transport is not so difficult to those people who bring their goods here from a far distant places to be sold. (See Map 2).

3 [4]On the southern side of the market is a school and a church near by and there is also a trading centre. [5]Villages are all around the market where great quantities of produce come for sale at the market. [6]The people in this area grow tobacco, groundnuts, maize, sugar-cane, peas and beans. [7]The fertility of the soil around this area has favoured the growing of these crops. [8]A branch of ADMARC was just set up last year to reduce the problems which the people had to sell their groundnuts and maize. (See Map 2).

4 [9]The mean annual rainfall ranges from about 35-40 inches per year. [10]We getthe heaviest rain in the months of January and February.

[11]The climate around the area is fair and suitable for the growing of these crops. [12]It is warm during the rainy
5 season and the mean temperature is around 30 C, and varies according to seasons.

[13]The market is just developing now. [14]At first it was a little market and it was less important. [15]But now it is
6 expanding year after year because of the streaming numbers of people who come to sell their produce. (See Map 3)

[16]The market has two gates (See Map 3) one on the southern
7 side and the other on the northern side. [17]The southern gate is the entrance and the other one is the exit.

[18]People around the area bring their different crops to
8 this market to sell. [19]Such crops are as follows: vegetables, sugar-cane, potatoes, onions, peas, beans, pumpkins, and millet. [20]Dried fish from Salima are also sold in this market, and, so say the fact, people are very fond of buying these chambo fish.

[21]The people have dimbas where sugar cane is cultivated
9 and also a few vegetables are grown.

[22]There is also a butcher where cattle and goats are slaughtered on every Wednesdays and Saturdays. [23]Before any
10 meat is sold it is first of all examined by a person from Vetrinary Department to check if the meat is worth selling here.

[24]Before a person brings goods into the market he has first of all to buy an admission ticket so that he has
11 the right to sell his/her goods freely. [25]The admission fee is three tambala per day.

[26]I have observed for a long time and I have found
12 that only sugar-cane is always found in large quantities. [27]I think its because it is also produced in large number.

[28]The people around the area find the market to be more beneficial in that it is a source of income to them,
13 and they also come to make friendships with other people from different **areas** or districts.

FIGURE 2 EXAMPLE OF STUDENT WRITING : SECOND YEAR TERTIARY

What are the main barriers facing the educationalist in attempting to bring about changes in attitudes towards modernisation in this society?

[1]Malawi is a developing society. [2]A developing society is, by definition, the society in which changes in attitudes
1 are intended. [3]But in changing attitudes there are many barriers to be reckoned with.

[4]In a traditional society most attitudes bind a group of people together, and each member feels insecured when he tries to change his attitudes. [5]It is then difficult to change an individual's attitudes unless one influences the leader first. [6]The theory of consonence will suggest that any attitides that are similar to those that already exist in the society will be accepted very quickly. [7]But the theory of dissonance suggests that if two opposing attitudes are introduced they will result into one of them being dropped out. [8]In schools there could be a conflict between the scientific and traditional way of solving problems. [9]The pupils may in most cases accept the scientific values just for the sake of passing
2 examinations, but when they go to their homes they return to the traditional ways. [10]In the traditional society competition between individuals does not exist. [11]Everyone would like to be at the same level as everyone else. [12]This brings a problem among the "Achikumbe". [13]The child who has been to school may know good methods of agriculture, but will not use them or put them to practice because he is afraid if he proves a success he may be bewitched, or this may displease his neighbours, who may also think their poor crops are a result of the medicine from the man with good and healthy crops.

[14]To convince the traditional man in the traditional society to come to believe in the power of fertilizer is a problem. [15]Many men in the traditional society will not send their children to school because they do not want their children to change after being exposed to institutions. [16]Parents are in most cases the major barrier to modernisation for it is them who have the whole influence on their children. [17]Talking
3 about agriculture, people will see no connection between hard work in the fields and economic development. [18]When the distance between the workers and advisers is so great it is very difficult to change the ordinary farmer's way of farming. [19]All these are problems that the educationalist will face in his attempt to bring about changes in attitudes towards modernisation.

2. English in Malawi

It might be useful at this point to present a brief survey of the language situation in Malawi, in order to suggest an explanation for the failure of the schools to produce good expository writers, and also to emphasise how important it is that such a failure should be rectified.

2.1. The role of English as a second language. Malawi is a small and relatively insignificant country, but in its use of English it is representative of other African anglophone countries, such as Zambia, Kenya, Ghana or Swaziland.

English was introduced into Malawi as a result of British rule, but it has retained its importance long after independence because of the absence of any one native language sufficiently dominant to supersede it. According to the 1966 Census, Chichewa is spoken as a mother-tongue by 50.2% of the population, and as a lingua franca by a further 26.4%, but is nevertheless a regional rather than a national language, occurring predominantly in the centre of the country. Although Chichewa was made an official language, jointly with English, in September 1968, it would probably be politically too risky for it to replace English, as the majority of people in the Northern Region (which still supplies much of the intelligentsia) speak Tumbuka, and large Lomwe and Yao minorities exist in the South. Thus, as in a number of other African countries, English is likely to be retained in the foreseeable future as the only language acceptable to all tribal groups in the country.

Although only 5.8% of the population claim to understand it (1966 Census) English is of key importance in the administrative

commercial and educational life of the country. English is the working language of Parliament, the law courts, the civil service, and commercial companies. It is the medium of instruction for the final two years of the primary school, the secondary school, and all tertiary level institutions. Several hours a day of English are transmitted by the Malawi Broadcasting Service, and the only national daily newspaper is published in English. Command of English is vital for anyone wishing to become a member of the educated élite class, the class who have been able to escape from the poverty and limited horizons of village life and enjoy the fruits of power. As the next generation passes through the educational system and this élite becomes larger, the position of English is likely to strengthen, and the necessity and demand for greater competence in it will increase.

2.2 English within the school system. A complete period of schooling lasts for twelve years, eight primary and four secondary, culminating in the MCE Examination (Malawi Certificate of Education). English is taught as a subject from the very beginning, gradually replacing Chichewa as the medium of instruction from Standards 6 to 8. By the time the pupil enters secondary school all subjects are taught through English, though it also continues as a school subject in its own right for eight thirty-five minute periods each week. An MCE pass in English is necessary for entry to the university and most other post-secondary institutions.

An unusual stress on formal grammar in the schools ensures that the emphasis in the English class is geared more

toward writing than is the case in most African countries with a similar background. This results from personal intervention by the Life President Dr. Hastings Banda, who, concerned at the falling standard of English, stated in a speech at the Kwacha Cultural Centre, Blantyre, in 1968 that "if English is to be learned properly, Grammar and Composition have to be taught". Thus, even at primary level pupils are expected to parse sentences in the traditional manner, and all three external school examinations contain a paper to test skill in continuous writing.

2.2.1 The composition paper in the MCE Examination. As the writing which led up to this investigation was that of first year university students, it is not necessary to consider the primary and secondary school syllabuses in detail. It might, however, be useful to outline the requirements of MCE Paper 2 (Composition), as this will indicate what is expected of a secondary school leaver who intends to go on with his studies at tertiary level.

The composition paper is divided into two sections, one of which contains a compulsory letter, and the candidate is required to write two essays, one from each section. Since the introduction of the MCE Examination in 1972, considerable help has been given in the subject-matter and vocabulary of the essays, on the grounds that the examination is a test of ability in English rather than originality of thought. The official objectives of the paper, according to the Ministry of Education (1972), are as follows:

The candidate should be able to

1. Read and quickly ascertain the requirements stated in the question paper, and the meaning and requirements of the topics set in it.
2. Make a quick choice of one topic in each of the two sections of the paper from the given number of alternative topics according to his personal preference.
3. Decide how to use his own ideas on the topic chosen, or those offered as subject-matter in note form, or proportions of both in combination. (Ideas emanating from the student must be relevant to the topic, the purpose of writing and the intended recipient)
4. Decide on the order in which he will present items of his own, and/or the given subject-matter.
5. Choose the form of language and presentation appropriate to the topic, its subject-matter, the purpose of writing and the intended recipient.
6. Demonstrate, in writing, his selection of legible, correctly spelt, grammatically correct sentences, appropriately punctuated in accordance with accepted sentence patterns, and without a monotonous sequence of any single pattern.
7. Produce, on each of the topics chosen, the equivalent of between $1\frac{1}{2}$ and 2 sides of "Exercise" ruled foolscap paper of the kind of writing described, and within such limits to write at least 300 and not more than 600 words ; and
8. perform the operation described above within one hour and thirty minutes.

It is clear from the above that some importance is attached to organisation (Point 4), although there is a greater stress on correct sentences (Point 6). There is no specific mention of paragraphing, cohesion between sentences (and groups of sentences) or overall coherence.

Some indication as to general levels of performance can be found in the examiners' reports. It is significant that although ability in paragraphing is not specifically demanded, its lack in performance is noted. The first MCE Chief Examiner's Report (1972) commended spelling, hand-writing and general presentation, but complained that paragraphing was poor or non-existent, while the 1973 Report stressed :

> Monotony of sentence patterns, indifferent paragraphing by some students and uncontrolled sentences remain, and there was little sign of improvement in these aspects.

The written work of first year students at Chancellor College also confirmed this judgment. While the few more able students could produce work of a high standard, the majority, and in particular those studying for the Diploma in Education, had difficulty in expressing themselves clearly and in an orderly manner. In some subject areas in the sciences essay writing was no longer required, it having been substituted by objective testing and short paragraph writing, but in most arts and social science subjects it was still felt that written expression was a necessary skill for a university graduate. Whatever the subject area, however, there was a growing feeling among the members of the college staff that something not easy to define was wrong with much of the students'

written work. The reasons for this failure are not far to seek : in spite of the stated objectives, teachers in the schools have been concentrating on sentence production to the almost complete exclusion of the continuous text.

2.2.2 The failure of the schools. It is not difficult to understand the local and immediate cause if we consider the special problems of the teacher in Malawi, and indeed in most developing countries. At primary level most teachers have themselves received very little education, many of the older "T3" teachers having only completed two years of secondary school before their professional training. At secondary level most are still diplomates rather than graduates, and, however efficient at their job some of these teachers may be, they are usually recruited from those who have passed the MCE, or the old School Certificate, at the lowest level. In no case have they anything like a perfect control of English themselves, and many live and work at isolated schools remote from the mainstream of new ideas in education or in ELT. Such teachers have need of a clear and explicitly laid out syllabus and a full and comprehensive course to follow, and this they do not have, at least at secondary level.

The effective syllabus of the first two years of secondary school up to the JCE Examination (Junior Certificate of Education) is a set of two textbooks especially written for Malawi and recommended to schools by the Ministry of Education : McAdam's Foundation Secondary English Books I and II. This lays heavy stress on grammatical structures, but does not teach

continuous writing at all, there being only an exercise in essay writing at the end of each unit, with little or no help given to the teacher as to how to present it to the pupil or to the pupil in carrying it out. A syllabus for the JCE was published by the Ministry of Education in 1973, but this confines itself to giving hints on methodology and recommending textbooks. It neither states terminal objectives nor offers any materials to assist in attaining such objectives. Suggestions for teaching writing include the combination of sentences, re-ordering jumbled sentences, and, at a later stage, situational composition, but there is no mention of paragraphing or cohesion. It is stated, however, that pupils at this stage should be able to arrange their ideas in logical order.

The English of the two final secondary years is based on two books from a course originally written for West Africa : Ogundipe and Tregidgo's Practical English Books 3 and 4. Although an improvement on the previously mentioned course-books, Ogundipe and Tregidgo also stress grammatical structures at the expense of essay writing. Thus, throughout the secondary school the teacher, who may not feel very secure in his or her own fluency and competence in English, is left to organise the essay writing component of his course with very little support from either the textbook or the official Ministry of Education syllabus. This is surprising, bearing in mind that the President himself has stressed the importance of both grammar and composition : the grammar element of the syllabus has been strengthened at both primary and secondary level, but composition has been effectively ignored.

3. Wider Influences

The educational system in Malawi, of course, does not exist within a vacuum, but is influenced by trends in educational theory evident in the world outside. As far as language teaching is concerned, these trends are strongly affected by current ideas in linguistics and psychology, or rather by ideas which have been current in the recent past. By the fact that English is a second language in Malawi, teaching has also been influenced by ideas about mother-tongue teaching recently current in Britain, with its stress on literature and creative writing.

3.1 Linguistics and psychology. In Malawi, English teaching has been influenced by the Direct Method (See Mackey 1965) and by the Audio-Lingual Method (See Rivers 1964). The Direct Method has been on the educational scene at least since the end of the last century, and is based on the idea that the second language learner will learn most effectively if he follows the path taken by the first language learner (that is, as far as this is possible in the artificial conditions of the classroom). It is thus a predominantly oral approach. According to Mackey (1965 : 151-152), it is characterised by such features as the use of everyday vocabulary and structures, the teaching of grammar orally and by situation, the use of visual aids, extensive listening and imitation until forms become automatic, the first presentation of all reading material orally, and the doing of most of the work inside the classroom rather than by private study.

The Audio-Lingual Method shares some features with the Direct Method, but is based on the more explicit theories of early Structural Linguistics and Behavioural Psychology. It was first worked out in the United States during the Second World War, when it became necessary to train military personnel to fluently speak a number of unusual languages, such as Japanese and Tagalog, within a limited period of time. Special schools were set up as part of the Army Specialised Training Programme (ASTP), and their courses were organised by linguists and anthropologists, with such well-known figures as Bloomfield, Boas and Sapir as advisers. As the goal of these courses was oral fluency, the spoken language predominated, with little or no reading, writing or grammar, and they mostly consisted of the imitation of informants, followed by drills supervised by linguists. The unprecedented success of the ASTP led to watered-down versions of its courses being used after the War in the foreign language programmes of schools and colleges throughout America, and to the dissemination of certain features of the method throughout the English-speaking world.

As far as Malawi is concerned, neither the Direct Method nor the Audio-Lingual Method has ever been used in the schools in its entirety, but certain elements present in one or both of them have helped to produce a situation in which the teaching of skill in expository writing has been neglected. These elements can be placed under two headings, the belief in the primacy of speech, and the study of the sentence as the highest grammatical unit.

3.1.1 The primacy of speech over writing. Because the child learns to speak before he learns to write, the Direct Method concentrates on oral language. The primacy of speech over writing was also one of the basic assumptions of the structural linguists. Bloomfield wrote that "writing is not language, but merely a way of recording language by means of visible marks" (Bloomfield 1933 : 21), and that "for the linguist, writing is, except for certain matters of detail, merely an external device like the use of the phonograph. . ." (Bloomfield 1933 : 282). In "A Linguistic Guide to Language Teaching", Moulton (1966) wrote detailed chapters entitled "Sounds", "Sentences", "Words" and "Meaning", but his final chapter, "Writing", contained only orthographic information on different types of writing systems.

Wilkins (1974) points out that this attitude has resulted in the allocation of most of the teaching time in schools to oral work, so that nowadays second language learners tend to be much more competent at speaking than at writing, the opposite having been the case when the grammar-translation method was fashionable. Teachers believe (though this attitude is admittedly now changing) that if the learner can speak correctly he will be able to make the transfer to writing without any difficulty. This is clearly a fallacy, as many mother-tongue speakers themselves never achieve competence in writing, even after years of practice at school.

3.1.2 The sentence as the highest grammatical unit. Bloomfield's famous statement that "each sentence is an independent linguistic form, not included by virtue of any

grammatical construction in any larger linguistic form" (Bloomfield 1933 : 170) effectively blocked the study of text for the next three decades or more. It is true that some structuralists, like Zellig Harris (1952) and Kenneth Pike (1955-1960) looked beyond the sentence, the latter attempting to incorporate levels above the sentence into his "unified theory of human behaviour", but the mainstream concentrated its attention on phonology, morphology and syntax.

The result in the schools was that considerable attention was given to the teaching and drilling of structures, without looking at sentences within the context of other sentences. Unfortunately, traditional rhetoric was not a significant factor in the British system to act as a counter-balance to such a sentence-based approach, an approach which may have been at the time linguistically valid, but was in many ways pedagogically disastrous.

3.2 Literature and creative writing. Traditionally, English teachers of UK origin in the secondary schools in Malawi have been graduates of English with a training in Literature rather than Linguistics, and there has consequently been a bias towards literature in their teaching. This has resulted in much of the writing in the English class being related to the "literature" currently being read, in the form of book reports, personal reactions to the texts, or standard School Certificate-type literature essays. Such a bias reflects the attitude still held by many teachers that second language teaching is little different from mother-tongue teaching.

In the 1970s, emphasis began to shift from British and American towards African literature, with an increasing interest in "creative" writing, especially in the English Department of the University. This resulted in an over-emphasis on the Expressive/Poetic mode, and a neglect of the Transactional mode (Britton 1975 : See 6.1 this chapter). Although such writing should be encouraged, especially in present-day anglophone Africa, where there is an ever-growing interest in the production of novels, plays, poetry and short stories (well-illustrated by the success of the Heinemann African Writers Series), this has its dangers. Lack of constructive criticism caused by over-anxiety to encourage individual creativity has resulted in much pretentious and badly-written student work. The student who receives praise for a poorly organised and grammatically incorrect short story or piece of personal writing will understandably not be very willing to strive to achieve a higher standard in expository writing, which presents very different problems, not always recognised by teacher or student.

4. A possible answer : English for Special Purposes (ESP)

The latest arrival on the ELT scene in Africa is English for Special Purposes (ESP). Perhaps to the rest of the world this is no longer new, but its extension to countries where English is used as a second rather than as a foreign language is recent. Although practitioners of ESP have up to now concerned themselves mainly with reading rather than writing skills, some teachers feel that an answer may be found here, and it has been suggested that ESP-type courses

should be set up for first year students at Chancellor College, probably in the areas of the Physical Sciences, Social Sciences and Law.

Without going into all the drawbacks and disadvantages of ESP, which have been briefly but adequately covered by Abbot (1978), two problems should be pointed out. First of all, where English is used as a second language, the aims of ESP are too narrow. The student needs to be taught to write not only with the immediate goal of learning science, or mathematics, or geography, through English, but also with the goal of functioning adequately after his studies in a society which, for the educated person at least, is partly English-speaking. If the student subsequently becomes a government or company official he will need to write letters, reports, memoranda and minutes in English, if he becomes a teacher he may wish to write an article for his professional journal, and in any case as a private citizen he will need to write numerous letters of one kind or another in formal or even official language. A narrow "English for Chemistry" or "English for Law" will be neither completely useful or acceptable.

Secondly, many African countries, including Malawi, have small secondary school populations, and as a result, by European standards at least, small universities. Lesotho, Botswana, Swaziland, Sierra Leone, Malawi and probably Tanzania, are not likely to support more than one small university or university college in the foreseeable future. Chancellor College, for example, has fewer than a thousand

students, spread over five years of study, and divided up between the usual arts and science subjects. Most students study for general degrees involving several different subjects, at least in the first two or three years, so it is impossible to divide all these students up into subject groups, partly because the resultant groups would be too small, but mainly because at the elementary stage at which they need such instruction no student can be said to have one "subject". In any case, it goes against the social and educational policy of the college to separate students according to subject and/or ability - English classes are deliberately mixed as much as possible.

Some kind of approach to the teaching of writing must therefore be found which is not closely geared to one subject area, and which can be applied to the type of writing the student might later be called upon to produce in his roles as adult worker and citizen. In this sense, it will complement "creative" writing, which is already catered for. This should be a "rhetorical" approach, which would, more specifically and formally than is done at present, teach the principles behind expository writing, so that they can be applied not only to the students' academic essays but also to any similar writing he may need to do in future.

5. General Aims

On looking into the problems of the Malawian undergraduate writer, it becomes more and more evident that little or nothing exists to help the teacher, or to help the learner to organise his work and paragraph it correctly beyond piecemeal prescriptions handed down for decades in textbooks

of rhetoric, and now in their modern counterparts, college composition handbooks. Such directives may or may not be based on reality : it is almost impossible without close analysis to say whether we find a particular statement, eg that a paragraph should normally begin with a topic sentence, acceptable because we ourselves at some point were taught such a rule, or because we have met large numbers of paragraphs structured in this way throughout a lifetime of reading. In any case, no clear definition yet exists of the topic sentence itself.

Nevertheless, rhetoric cannot, and should not, be ignored. As must be obvious by now, there is a major gap at discourse level in the teaching of English as a second language, and dismissing rhetoric as old-fashioned does not help us to fill this gap. An alternative course of action would be to take from rhetoric whatever is worthwhile, but first it is necessary to confirm or reject, by experiment or otherwise, the validity of its major prescriptions.

Following this path, it was decided to investigate certain areas of text organisation where the Malawian students seemed to have special difficulty, in an attempt to formalise what up to now has been piecemeal. The view is taken that whereas many of the prescriptions of traditional rhetoric are useful and productive in the teaching of writing skills, they should not be presented to students sheerly on a "folklore" basis, but should if possible have some experimental backing. In any case, teaching should be more effective if a given prescription is seen by the teacher to be part of a

larger whole, if the theory behind the prescription is made explicit to him. The following three areas are central to the investigation :

(i) The organisation of text from sentence-level upwards

In the words of Pike and Pike (1977 : 1), "Purposive behaviour, including language, comes in 'chunks' or 'constituents'". Sentences do not simply follow each other with adjacent cohesive links - chunks of text are linked to chunks of text, and these smaller chunks combine together into larger ones. An attempt will be made to discover some of the principles behind this chunking, and its relationship to paragraphing. Viewing the sentence itself as a lower-level information block, optimum sentence length will be investigated.

(ii) The paragraph Some light will be thrown on the traditional concept of the paragraph, including its reality or otherwise as a discourse unit, its structure, the topic sentence and the bridge sentence.

(iii) Signals of text chunking Both implicit and explicit signals are used by the writer to indicate movement from one information block to another, one characteristic of the second language learner as a writer being his failure to use these signals adequately. An investigation will be made into the identity of some of these signals, and their relative effectiveness.

6. Some definitions

In an area which remains largely unformalised, terms tend to be used vaguely and often inter-changeably. Before

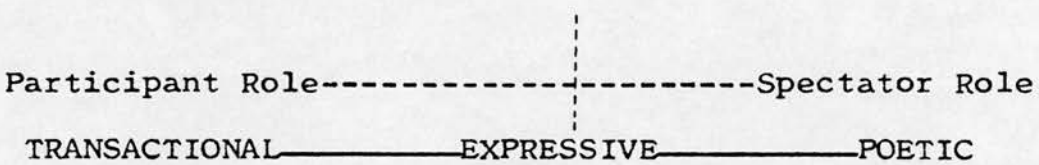
beginning it is therefore necessary to define some of those terms which will appear in the experimental part of the succeeding pages.

6.1 Expository prose As the purpose of this work is primarily practical, the writing categories of Britton et al (1975) have been used, categories produced as part of the Schools Council Project on the Written Language of 11 to 15 Year Olds (1966 - 71), and used for the categorisation of actual samples of writing.

Britton identified three main functional types - Transactional, Expressive and Poetic - depending on the role of the writer either as participant or as spectator. (Figure 3)

BRITTON'S THREE FUNCTIONAL TYPES

FIGURE 3



(From Britton 1975 : 81)

In the participant role, the writer uses language to take part in the activities of the world ("operating in actuality via representation" (Britton 1975 : 80)). In the spectator role the writer uses language simply to view what is going on in the world, seeking in no way to affect or influence these happenings ("working on the representation without seeking outcomes in actuality" (Britton 1975 : 80)). Expressive language occupies the mid-point of what is in fact a cline,

and it is represented by the ordinary, everyday language of the home and close relationships. Examples in writing would include personal letters, diaries and some popular newspaper articles. As the writer needs to write for a wider and more public audience, he moves along the cline either towards the poetic end where form becomes central, resulting in language as art, or towards the transactional end where function becomes central, resulting in language for informing, persuading, presenting opinion, etc.

Britton et al divide the Transactional into two sub-categories, which they term Conative and Informative, and it is the latter which can be used to define expository prose as it is considered here. Britton's seven types of informative writing are as follows, in ascending order of generality :

(i) Record An eye-witness account of what is immediately present.

(ii) Report A description of observed past events.

(iii) Generalised narrative, or descriptive information Similar to the above, but more generalised, although without abstraction. It could be an account of events, based on a time sequence, or a description of a place based on spatial patterning.

(iv) Analogic : low level of generalisation Loosely organised generalisations with no attempt to present an overall view.

(v) Analogic Logically or hierarchically related generalisations.

(vi) Analogic/tautologic (speculative) The generalisations become the subject of the writing, with loosely diciplined open-ended speculation.

(vii) Tautologic This form is concerned with theory, and involves hypothesis-making and deductions from hypotheses. It is more organised and disciplined than analogic/tautologic.

The other aspect of the Transactional, the Conative, is sub-divided into the Regulative and the Persuasive forms. Although often considered as types of expository writing, these are excluded for the purposes of this study.

6.2 Discourse, text and passage Discourse is a general term which indicates a level of organisation above grammar (Coulthard, 1977 : 6) (The present writer takes the view that there is no grammar in the normally accepted sense above the sentence). Discourse is manifested as text. A text is brought into being in spoken or written form, and is defined by Halliday and Hasan as

> ...a passage of discourse which is coherent in these two regards: .... it is coherent with respect to the context of situation, and therefore consistent in register ; and it is coherent with respect to itself, and therefore cohesive.
>
> (Halliday and Hasan, 1976 : 23)

Thus, the terms "discourse" and "text" as used here are not equivalent to those suggested by Widdowson (1973). Widdowson's interpretation is illuminating, but it was felt that the distinction he makes between text and discourse was not relevant to this particular study, biased as it is towards

text analysis rather than discourse analysis (in the sense used by Widdowson).

Passage refers to an extract from a particular text used as the basis of an experiment or an analysis.

6.3 Paragraph and information block. As will become clear later, confusion has frequently arisen due to failure adequately to define the paragraph. Here it is considered as a physical unit, and refers to that part of a text marked by a new line and an indentation. (The 'orthographic paragraph' of Lackstrom, Selinker and Trimble (1973)). It depends on external as well as on internal features, including the idiosyncratic preferences of the writer.

The information block may or may not correspond to the paragraph. It is a chunk of text one sentence or more in length which hangs together as a unit on account of its internal relationships. It is similar in some respects to the 'conceptual paragraph' of Lackstrom, Selinker and Trimble (1973).

7. Overall plan and hypotheses

What follows can be divided into three parts, chapters 2 and 3, chapters 4 to 8 and chapters 9 and 10.

7.1 Chapters 2 and 3 These chapters will deal with the prescriptions of traditional rhetoric concerning paragraphing, together with the work of a number of American teachers of college composition who have attempted to explain paragraphing by constructing surface-level models. Traditional rhetoric, the basis from which much of the experimental work begins, is approached by surveying fifteen representative textbooks, one from each decade

between 1866 and 1978. The successors to these traditional writers, American teachers of college composition such as Frances Christensen and Alton Becker, will then be considered. These were active mainly in the 1960s. Like their forebears their motivation was pedagogical rather than theoretical.

Since then, interest in paragraphing has waned, although occasional helpful insights can be gained from such disciplines as psychology, philosophy and artificial intelligence. However, as the purpose of this work is to demonstrate that traditional prescriptions can be experimentally upheld, such areas will be reserved for consideration in some other place.

7.2 Chapters 4 to 8 These chapters are concerned with relevant experimental work. No attempt will be made to build up a water-tight model, but it is hoped to

(i) discover whether there is any kind of reality behind the traditional concept of the organic paragraph

(ii) discover whether there is any reality behind the traditional concepts of the topic sentence and the bridge sentence

(iii) present evidence for the traditional prescription that sentence length should be varied, and indicate one way in which this might be carried out

(iv) suggest which types of signals can best sign-post movements from one information block to another within the text.

Five hypotheses will be tested within four experiments, and some tentative theories will be suggested on the basis of the analysis of chosen texts.

7.3 Chapters 9 and 10. While the main aim of this work is to provide a theoretical justification for the utilisation of traditional rhetoric in the teaching of writing skills, some more direct pedagogical applications arising from the experimental work will become evident. Chapter 9 will present a theoretical summary, and Chapter 10 will conclude the investigation by relating the findings to the classroom.

CHAPTER 2

THE LEGACY OF TRADITIONAL RHETORIC

1. Introduction

A serious weakness among applied linguists and educators working in a period of rapidly changing ideas is their neglect of the insights of earlier writers. Modern theorists are building upon ideas already expressed by earlier workers even if they are not always aware of this. Very little in the world is in fact new.

In this chapter, the subject under examination will be the theory of composition writing, and in particular the paragraph, as presented by some of the prescriptive textbook writers of the last hundred years, those writers who have been, and are, following the tradition of rhetoric first formulated by the Classical Greeks. This tradition has had its effect on all of us through the transmission of normative rules, often piecemeal and unorganised, in the secondary schools, and is the source of most of the commonly held views on writing prevalent among educated laymen. A number of currently fashionable theories now percolating into the educational system, in particular those to do with cohesion, coherence and speech act theory, have their roots in the ideas of these writers.

2. Historical background

Before going on to consider individual texts in detail, it is necessary briefly to summarise the historical development of traditional rhetoric, leading to a concentration of interest on the paragraph in the twentieth century. The lack of emphasis in the teaching of writing skills in Britain as compared with the United States will be touched on.

2.1 Aristotelian rhetoric was a basic discipline in Western education for over 1500 years, and together with logic and grammar was a member of the trivium studied in schools and universities in medieval times. The growth of democracy in Classical Athens was a spur to the development of various theories of rhetoric, their aim being to discover the most effective way of convincing an audience of common people in the market place of what the orator considered to be true and just. Thus Classical Rhetoric was concerned primarily with oral discourse, and with persuasion rather than with exposition, description or narration.

Aristotle laid down five "arts" in the production of an effective speech. These were invention (inventio), arrangement (dispositio), style (elocutio), memory (memoria), and delivery (pronuntiatio). Memory and delivery are exclusively the concern of oral discourse, so will be ignored here, but the first three "arts" are equally applicable to the written form. The first, invention, was concerned with the discovery of relevant arguments. These arguments might be either extrinsic, ie argumentation based on evidence, such as documents, experimental observation or eye-witness accounts,

or intrinsic, ie argumentation based on emotional and ethical appeal. Arrangement, the second art, was concerned with how the argument was to be developed, and the third, style, was the technique of framing effective sentences. As time passed, emphasis moved away from clarity and towards elegance, so that eventually style became an ornament and an end in itself.

As far as the subject of this work is concerned, the second art, arrangement, is the most central. Various patterns of argument were suggested. For example, the Ciceronian plan had six parts, as follows :

(i) Introduction (exordium)

(ii) Narrative (narratio) - exposition of the problem's history

(iii) Proposition (divisio) - outline of the steps in the argument

(iv) Demonstration (confirmatio) - proof

(v) Refutation (confutatio) - refutation of alternative propositions

(vi) Peroration (peroratio) - conclusion

(From Corbett 1971)

It was the duty of the orator, given this guide, to distribute his material between these six stages, deciding where to emphasise, expand, or even re-order, depending on his material and his audience.

2.2 The last hundred years Unlike the other members of the trivium, rhetoric has changed little to the present day, and

the influence of Aristotle can still be clearly seen in many of the handbooks on college composition even now being published in the United States. Developments in science since the seventeenth century, however, have led to more emphasis being placed on types of prose other than the persuasive, and in particular on exposition. Another major change was that oral delivery, once dominant, gradually became separated off as elocution, while writing remained as the central art of composition. According to Rodgers (1965) this division had come into existence in the United States by the 1880s.

From that time onwards, the way lay open for a formulation of a theory of the paragraph. Classical Rhetoric, as we have seen, concerned itself with overall planning (arrangement) and the sentence (style). Probably because the end product was oral rather than written, the level between the broad macro-structure of the text and the sentence was ignored, intermediate structuring being left to the orator. This is the level that is most difficult for the modern student.

Although the prescriptive teaching of composition writing is a continuing tradition in the United States, it has never been of great importance in Britain (with the possible exception of Scotland). Most British books on writing produced in the last hundred years have had the average educated person rather than the student as their target, and have concentrated on idiom and usage at and below sentence level. Indeed, Albutt, in one of the earliest works on writing for scientists, said :

> Chapter, and even paragraph, I say, we may dispense with, but the sentence we cannot ignore.
>
> (Albutt 1904 : 61)

Quiller-Couch (1916) one of the most influential figures in this area, concentrated on style and paid no attention at all to organisation, and Fowler and Fowler (1930) in their much-read manual "The King's English" similarly, explicitly stating in the preface that their principle had been "to pass by all rules, of whatever absolute importance, that are shown by observation to be seldom, or never, broken." Such writers as Whitten (1939) and Vallins (1951) completely ignored anything above the sentence. A book still widely used, Partridge's "Concise Usage and Abusage" (1954) directly quotes Alexander Bain (1866) (see below) in a short section on the paragraph, but nevertheless seems to suggest that paragraphing and higher level organisation are natural processes, and an inevitable result of the intrinsic ordering of the material. Gowers (rev. 1973), an equally influential writer, gives a little advice, but warns that "the subject does not admit of precise guidance". (258)

It is tempting to hypothesise that the continual dominance of the oral tradition in rhetoric at the beginning of the century, and the influence of structural linguistics from the 1930s onwards are the main reasons for the absence of the prescriptive teaching of composition writing in Britain. This, however, would still leave unanswered the question as to why the situation in Britain was different from America. The key factor is probably social. The British tradition has

never seriously encompassed the teaching of writing skills because, with an élitist educational system, it has always been assumed that the educated person would absorb the ability to write as a result of wide reading and a cultured background. The ethnic diversity of the United States, on the other hand, together with a more democratic attitude towards education, has meant the presence in the schools and even in the universities of large numbers of working-class students with recent immigrant backgrounds. Thus, the teaching of writing has always seemed more necessary in the United States than in Britain. Now, however, we are beginning to see a conscious recognition growing in Britain that changing circumstances demand a different approach.

2.3 A theory of the paragraph Although most of the textbooks on rhetoric produced in the last hundred years have been American, the first writer to develop a theory of the paragraph was in fact a Scot, Alexander Bain (See Rodgers 1965), whose "English Composition and Rhetoric" was first published in 1866. Bain was Professor of Logic at Aberdeen University between 1860 and 1880, and his interest in rhetoric sprang from disquiet at the poor standard of writing shown by the students there.

Rodgers suggests that the reason why these student writers were having such difficulty at paragraph level was "sentential simplification", ie. that the average sentence length in English prose was in process of drastic reduction. Flesch (1949) quotes Sherman (1893) to the effect that whereas the average Elizabethan written sentence had had an average

length of 45 words, the average Victorian sentence had been reduced to about 29 words. One result of this was that the paragraph, staying at approximately the same length, now contained a greater number of sentences. Lewis (1894) wrote :

> We may safely conclude that the paragraph of today contains at least twice as many sentences as did that of Ascham's day. Indeed, if we accept Macauley's 'England' as a present-day norm, the past increase in the number of sentences per paragraph will be far more than one hundred percent in 300 years.
>
> (Lewis, 1894 : 42)

Coherence within the sentence is easier to maintain than coherence between sentences, and thus Bain was finding among his students an ever-increasing tendency towards a breakdown in continuity within the paragraph. The result was Bain's emphasis on unity and coherence in paragraph writing, principles which, first recommended by Bain, have been repeated again and again in prescriptive textbooks up to our own time.

3. Survey of representative texts

We will now proceed to examine a selection of fifteen textbooks of composition and rhetoric published between 1866 and 1978, to see what contributions traditional rhetoric has made to the theory of the paragraph and text organisation. All but one of these texts, the first, (Bain 1866), are American, and most have first year college students as their target audiences, with the exception of Canby (1909) and Genung and Hansen (1915), which were intended for secondary school pupils. A full list of the texts can be found in Appendix 1.

The most noticeable feature evident from these texts is the changelessness of the principles they advocate and the terminology they use. As we shall see later, it was not until the 1960s that any new ideas from other disciplines, notably linguistics, began to find their way through into such books.

3.1 Definition of the paragraph From Tables 1 and 2 it can be seen that the most common elements of the definition of the paragraph are that it is "a group or collection of sentences" which develops "one single topic", and that this topic presents "a unit of thought". This appropriately re-echoes the two earliest writers' definitions, ie, Bain's, that the paragraph is "a collection or series of sentences with unity of purpose" (Bain, 1866 : 91) and Genung's, that it is "a connected series of sentences constituting the development of a single topic" (Genung 1886 : 193).

Such definitions are clearly related to traditional definitions of the sentence. Fries (1952 : 9) points out that the typical school grammar definition of the sentence as "a group of words expressing a complete thought" antedates Priscian. Bain, in the 1887 edition of his book, makes the supposed similarity between the structure of the sentence and the paragraph explicit, when he states :

TABLE 1

DEFINITIONS OF THE PARAGRAPH 1866-1978

WRITER	DEFINITION
BAIN 1866	... a collection or series of sentences with unity of purpose (91)

WRITER	DEFINITION
GENUNG 1886	A paragraph is a connected series of sentences constituting the development of a single topic. (193)
BALDWIN 1902	A paragraph is a part, which, during the process of composition, has defined itself as one distinct stage in the progress of the essay. It is a unit, but a component unit. (11)
CANBY 1909	A paragraph is a unit, because it consists of a single thought developed as far as need be for complete clearness. (76)
SCOTT & DENNEY 1911	While the composition is being written, each topic in the plan - each fact or group of related facts - grows into a group of sentences that belong together. These groups of related sentences are called paragraphs. (44)
GENUNG & HANSEN 1915	No definition
SMART 1922	A paragraph is a group of sentences dealing with one unit of thought which forms a definite stage in the development of the subject of the theme. (27)
HULBERT & HULBERT 1929	A paragraph is a miniature composition which usually develops one topic of the outline. (34)
KIERZEK 1939	... a sentence or a group of sentences developing a single complete idea. (62) ... a unit of thought - or, more exactly, a unit of the expression and communication of thought - larger than the sentence and smaller than the section or the chapter. (62)

WRITER	DEFINITION
SHAW 1946	a paragraph is a group or bundle of sentences developing either one single topic or a specific part of a larger topic. (48)
BROOKS & WARREN 1950	Paragraph divisions signal to the reader that the division so set off constitutes a unit of thought. (290) A paragraph undertakes to discuss one topic, or one aspect of a topic. (291)
THOMPSON 1957	A paragraph is a group of sentences, or at times a single sentence, pertaining to one particular topic. (448)
JONES & FAULKNER 1961	... an independent clause or clause-group which functions as a separate portion of a written work and presents a particular idea or point that is a distinct unit of thought in the development of the subject of the whole work. (135)
KANE & PETERS 1966	... a structure of sentences unified by their common relation to a general conception - whether this be a feeling, an idea, an opinion - which we call the topic of the paragraph. (135)
WOHL 1978	... a sequence of sentences organised around a single idea; a group of related sentences, which, when logically and sequentially arranged, explain and develop a particular idea. (22)

TABLE 2

DEFINITIONS OF THE PARAGRAPH 1866-1978 : COMMON ELEMENTS

BAIN 1866	a collection or series of sentences			
GENUNG 1886	a connected series of sentences	a single topic		
BALDWIN 1902			a component unit	
CANBY 1909			a unit	a single thought
SCOTT & DENNEY 1911	group of related sentences	topic		
GENUNG & HANSEN 1915	N O D E F I N I T I O N			
SMART 1922	group of sentences		one unit of	thought
HULBERT & HULBERT 1929		one topic		
KIERZEK 1939	a group of sentences		a unit of thought	
SHAW 1946	a group or bundle of sentences	one single topic or a part of a larger topic		
BROOKS & WARREN 1950		one topic or one aspect of a topic	a unit of thought	
THOMPSON 1957	a group of sentences	one particular topic		
JONES & FAULKNER 1961	an independent clause or clause-group		a distinct unit of thought	
KANE & PETERS 1966	a structure of sentences	topic	unified	
WOHL 1978	a sequence of sentences			

> The internal arrangement comes under laws that are essentially the same as in the sentence, but on a greater scale.
>
> (Bain 1866 : 91)

and Genung develops this :

> The principle on which the plan of a paragraph is constructed may be regarded as an extension of the principle of sentence structure . . . .
> The same relation exists between sentences in the paragraph as between clauses in the sentence.
>
> (Genung 1886 : 199)

A few writers, however, viewed the matter from the opposite point of view, and compare the paragraph to the whole composition rather than to the sentence. For example, Hulbert and Hulbert (1929), who call the paragraph "a miniature composition", and Kane and Peters (1966), who point out that their definition of the paragraph as a "structure of sentences unified by their common relation to a general conception" could equally well be describing the essay, or even the whole book.

One writer, Canby (1909), is aware of the difficulty involved in using such a concept as "unit of thought" in a definition. He goes as far as to talk at length about the difference between "composition thoughts", "paragraph thoughts" and "sentence thoughts" (Canby, 1909 : 76-77). His example of a "composition thought" is "An ambitious boy should go to college". In order to develop this thought, he explains that it would need to be split up into a number of topics, each of which would be a separate "paragraph thought".

In contrast, the more concrete and specific, and therefore more limited, "If you do not hurry you will miss the train" he claims as a "sentence thought", it being almost impossible (he thinks) to develop and expand.

For all his words on the subject, Canby does not take us much nearer to the meaning of "a single thought", a particularly serious failure when one considers that all these books are practical textbooks. Similarly, "a single topic" is an unhelpful concept, though admittedly it would be possible for the student to build up his own intuitive understanding from the numerous examples found in most of these books.

Even the first part of the definition is not wholly satisfactory, in spite of its measurability. Few writers indicate that one sentence alone may stand as a paragraph, and in spite of stress on coherence in other places in the books the definitions do not usually show that any sort of linkage is necessary between successive sentences.

Familiar as all of these writers' definitions seem to most of us, not one stands up as a formal and testable definition within an explicit theory. An attempt will be made later in this work to come a little closer towards such a theory.

3.2 Bain's Six Rules Bain, described by Lewis (1894 : 29) as "the ablest writer on rhetoric since Aristotle", laid down the following six rules for the writing of a satisfactory paragraph :

I "The bearing of each sentence on what precedes shall be explicit and unmistakable".

II "When several consecutive sentences iterate or illustrate the same idea, they should, as far as possible, be formed alike. This may be called The Rule of Parallel Construction".

III "The opening sentence, unless so constructed as to be obviously preparatory, is expected to indicate with prominence the subject of the paragraph".

IV "A paragraph should be consecutive, or free from dislocation".

V "A paragraph should possess unity; which implies a definite purpose, and forbids digressions and irrelevant matter".

VI "As in the sentence, so in the paragraph, a due proportion should obtain between principal and subordinate statements".

With the exception of Rule II, all the above appear in one form or another in most textbooks of composition published since 1866. Indeed, although Rule II did not occur in the 1886 edition of Bain's classic, all six were recommended exactly as they stand by Grierson as late as 1944.

3.3 Genung's Three Principles John Genung, arguably the most influential early American rhetorician, reduced the six rules above to three in his "Practical Elements of Rhetoric" of 1886. These three principles were Unity, Continuity and Proportion.

Unity included Bain's Rules III and V. Genung explained :

> A paragraph is a distinct division of the discourse, related, indeed, to preceding and following, as a link in a larger chain, but complete in itself, and exhaustive of its topic. Its primary requisite, therefore, is unity, and this unity is subserved by choosing for each paragraph a determinate subject, to which all parts of the structure are related as constituting elements in its development.
>
> (Genung 1886 : 194)

He went on to modify Rule III by stating that although it was usual to present the subject of the paragraph in the first sentence, this was not strictly necessary. It could, on occasion, appear at the end, or even, for emphasis, at both beginning and end.

Continuity combined Rules I and IV

> .... all paragraphs should manifest a logical progress of thought, developing its suggestions of the subject from point to point, and without dislocations. Further, the bearing of one point on another should be clearly indicated throughout; and the topic should be brought to a complete and properly rounded conclusion.
>
> (Genung 1886 : 198-9)

Genung suggested that if possible it was better to do without explicit connectives to preserve continuity, but if this was too difficult, either because of the inexperience of the writer or the type of subject matter, he suggested using conjunctions or conjunctional phrases, demonstratives, or repetition.

Proportion was practically synonymous with Rule VI.

> on the principle that all statements should have bulk and prominence according to their importance, a due proportion needs to be maintained between principle and subordinate ideas in the paragraph. Every part should be so treated as to show for just what it naturally is, in rank, and in its relation to the whole.
>
> (Genung 1886 : 207)

He points out that if a subordinate idea is expanded too much it becomes a diversion, thus destroying the notion of unity.

3.3.1 Unity Genung's first principle, that of unity, has been accepted unaltered until the present time. (See Table 3). Some quotations will illustrate how static the theory of traditional rhetoric has been during the last eighty years in this respect.

> A paragraph has unity when it can be summed up readily in a single sentence.
>
> (Baldwin 1902 : 12)

> If a paragraph can be summed up in a single sentence, of which it is a true development, it has a unity.
>
> (Canby 1909 : 88)

> A good paragraph must be unified. Oneness of purpose is desireable ; extraneous detail must be eliminated.
>
> (Shaw 1946 : 50)

> A paragraph which has unity confines itself to the development of one central idea.
>
> (Jones & Faulkner 1961 : 98)

> .... another characteristic of a good paragraph is unity. A unified paragraph can be compared to a single cell, the fundamental structural unit of plant and animal life.
>
> (Wohl 1978 : 23)

TABLE 3

BAIN'S SIX RULES AND GENUNG'S THREE PRINCIPLES 1866 - 1978

BAIN 1866	III	V	I	IV	(II)	VI	
GENUNG 1886	Unity		Continuity		Proportion		
BALDWIN 1902	Unity		Coherence		Emphasis		
CANBY 1909	Unity		Coherence		Emphasis		
SCOTT & DENNEY 1911							
GENUNG & HANSEN 1915	Unity		Coherence		Emphasis		
SMART 1922							
HULBERT & HULBERT 1929	Unity		Coherence		Emphasis		
KIERZEK 1939	Unity		Order		Proportion		
SHAW 1946	Unity		Order		Proportion		
BROOKS & WARREN 1950	Unity		Coherence		Emphasis		
THOMPSON 1957	Unity		Coherence		Emphasis		
JONES & FAULKNER 1961	Unity		Coherence		Emphasis		Completeness
KANE & PETERS 1966	Unity		Coherence * Flow **				
WOHL 1978	Unity		Coherence				

* = unity of thought

** = unity of form

3.3.2 Coherence (See Table 3) What Genung labelled "Continuity" was renamed coherence by Wendell (1893). Shaw (1946) seems to have used Genung as his source, and prefers to talk about "order" rather than "coherence", while Kane and Peters (1966) see "coherence" as a sub-division of "unity" (as "unity of thought", together with "flow", which they see as "unity of form"). Otherwise, however, this second principle appears again and again, as can be seen by the following quotations :

> Coherence in a paragraph demands, first, a logical sequence of sentences.
>
> (Baldwin 1902 : 12)

> In an ideal composition, each paragraph leads up naturally to the next, like a link in a chain. So, too, the sentences in a well-made paragraph form a chain. The current of thought should be absolutely continuous from beginning to end : one unbroken progress. In other words, the paragraph should have coherence.
>
> (Genung and Hansen 1915 : 27)

> .... throughout the theme the thought should flow from sentence to sentence, from paragraph to paragraph, without a break.
>
> (Hulbert & Hulbert 1929 : 37)

> The thoughts within paragraphs should make orderly and clear progress, and there should be clear passage from one paragraph to another.
>
> (Shaw 1946 : 50)

> A paragraph has coherence when its central idea is developed in a logical, orderly fashion.
>
> (Jones & Faulkner 1961 : 99)

Most of these writers go into some detail over the use of explicit connectors to aid the production of coherent text. Genung and Hansen advise :

> Let one of the final tests which you apply to each paragraph of your written work be the effectiveness of your connecting links.
>
> (Genung & Hansen 1915 : 208)

According to Shaw :

> Coherence depends upon clarity of thought, but the secret of coherence lies in the use of connectives, transitional expressions.
>
> (Shaw 1946 : 61)

Bain recommended that the connections between sentences should always be made explicit, though later writers, such as Genung (1893), Kierzek (1939) and Kane and Peters (1966) suggest that it should be possible to write coherent prose without using too many connectors, and that the use of explicit signals alone will not bring order to a badly organised text.

We should not, therefore, make the mistake of believing that the at present fashionable stress on cohesion is a modern innovation. Bain's first rule is quite clear, and in the second edition of his classic he lists various types of connectors that can be used in some detail. Among these he includes conjunctions (cumulative, adversative and illative), demonstrative phrases, and a group that he labels as "other words and phrases", which includes expressions to state opposition, eg on the contrary, to return after a digression, eg to resume, to sum up, eg in short, and transitions, eg up to this point. Genung mentions some of

the same classes of connectors, as does Baldwin (1902).

Canby (1909) comments :

> There must be signposts, even within the paragraph. In the whole composition, transitional sentences and transitional paragraphs perform this service. Within the paragraph it is the word and the phrase upon which we must depend. The invaluable 'however', 'but', 'also', 'nevertheless', 'furthermore', 'finally', the no less useful 'of course', 'on the contrary', 'but to repeat', 'on the other hand', 'to return', are the signs which point the way through the thought development of the paragraph.
>
> (Canby 1909 : 89)

3.3.3 Proportion (See Table 3) Genung's principle of proportion, which was more or less synonymous with Bain's sixth rule, was later influenced by Wendell's Theory of Mass (Wendell, 1893). Wendell believed that the beginning and ending of a paragraph were the most emphatic and eye-catching places, and thus most suitable for the location of important information. Baldwin (1902) combined Genung's Principle of Proportion and Wendell's Principle of Mass under a single heading - Emphasis - and this third Principle has been passed down through generations of textbooks. A few quotations will illustrate, again, how unchanging the conception of "emphasis" has been.

> Emphasis in a paragraph depends naturally somewhat upon proportion. You must develop most extensively that part of your thought which seems to be most important.
>
> (Canby 1909 : 89)

> (The paragraph) should be proportioned so as to give most space to its main ideas, and should be arranged so as to end on an important idea.
>
> (Hulbert & Hulbert 1929 : 35)

> A good paragraph is well-proportioned. If the thought of the paragraph is important, the paragraph should be so fully and completely developed that the reader can easily understand the significance of that thought ; if the paragraph discusses an idea, or group of related ideas of comparatively less importance, the proportion of the paragraphs should reveal the difference in weight.
>
> (Shaw 1946 : 50)

> Sentences expressing the central idea and the main sub-points should be given great prominence.
>
> (Thompson 1957 : 164)

> According to the principle of emphasis, the relative importance of an idea or fact should determine the amount of stress it receives.
>
> (Jones & Faulkner 1961 : 100)

3.4 The topic sentence It is not certain where the term 'topic sentence' first appeared, but its existence was suggested by Bain (1866) in his first rule, namely, that "the opening sentence .... is expected to indicate with prominence the subject of the paragraph." This limits it to first position only, a prescription later relaxed by Genung (1886), who suggested that the theme of the paragraph, while usually stated at the beginning, could sometimes occur at the end, or even at both beginning and end. Neither writer seems to have carefully studied authentic text, or they would have realized that an explicit statement of the theme is probably the exception rather than the rule, or at least that it is only normal in a very particular type of expository prose.

Canby (1909), however, does realize this, saying that in narrative and description there is often no topic sentence.

In exposition, though, it is usually at the beginning of the paragraph. He writes :

> It is placed there as the expression $(x + y)^4$ is placed at the head of an algebraic operation of which the completion is the expanded form.
>
> (Canby 1909 : 87)

Genung and Hansen (1915), writing for secondary school pupils, recommend :

> The topic sentence comes sometimes at the beginning of the paragraph, sometimes near the middle, sometimes at the end. In many cases, however, (for instance, if the writer wishes the reader to <u>feel</u> or <u>realize</u> something - say the beauty of a scene or the pathos and fun of an event -) there will probably be no need of trying to state definitely the substance of what the paragraph is to include. The main idea is not so likely to be expressed at some point as it is to be diffused as an influence through the whole.
>
> (Genung & Hansen 1915 : 19-20)

This prescription seems so general as to be useless for the learner. Christensen characterises the implied topic sentence well when he refers to it as "a sort of ectoplasmic ghost hovering over the paragraph" (Christensen 1965 : 144)

Later writers all suggest, with greater or lesser emphasis, that

1. Every paragraph should have a topic sentence, either expressed or implied. (An implied topic sentence is a sentence which could be used to sum up the main theme of the paragraph)
2. This topic sentence usually appears at the beginning, but sometimes also at the end, and is possible anywhere in the paragraph.

3. It should be clear and concise, and is normally declarative, though sometimes interrogative.

Some writers (eg Kane and Peters 1966) suggest that the relation of the topic sentence to its paragraph is the same as the relation of the subject of the sentence to the sentence as a whole. Sentences which develop the topic are thus considered to function in the same way as the predicate.

The hierarchical organisation of topic sentences, in which one topic sentence may dominate several paragraphs, each of which has its own topic sentence and perhaps layers of subordinate topic sentences, is only dealt with by Jones and Faulkner (1961), the only writers here to attempt a unified model for paragraph writing. Kane and Peters (1966), taking a more traditional approach, touch obliquely on this important aspect in their discussion of "compound paragraphs", ie paragraphs which develop more than one topic and therefore contain sub-topic sentences. Kane and Peters are feeling their way towards seeing text as hierarchical structure, but never express this idea overtly.

It is easy to criticise what is again and again a diffuse and meaningless treatment of the topic sentence. According to these writers, it can be present or absent, it can appear anywhere at all in the paragraph, and it can be recognised because it expresses "the central idea" (Jones & Faulkner, 1961), (Kierzek, 1939), or "the topic" (Smart, 1922) of the paragraph. Nowhere is it explained what, in linguistic terms, the topic sentence is, and the over-simplified picture is given of an essay as a string of paragraphs, each with its

one and only expressed or implied topic sentence. Perhaps this is the least successful feature of the traditional rhetorician's treatment of text structure.

It should, however, be remembered that these writers are in the business of instruction, not research. The topic sentence is so stressed as a component of every paragraph not because this is a necessary feature of all good writing, but because it is a useful lifeline for the inexperienced writer engaged in learning a practical skill. Many of these books use prolific examples and illustrations, and it is hoped that the learner will discover the reality of the topic sentence from studying these examples and writing paragraphs for himself following the patterns they offer.

3.5 Paragraph development Another major weakness of traditional prescriptive rhetoric has been its lack of detailed treatment at the level between the sentence and the paragraph. Development of the whole essay is usually considered, by such means as chronological arrangement, enumeration or cause and effect, but paragraph development is normally left to the student to work out for himself.

The only early writer to offer any guidance was Genung (1893), and indeed his formula for paragraph development was the most original contribution until the 1960s, when teachers of writing began to come under the influence of development in related disciplines. Genung's formula is as follows. He suggested that its different parts could be expanded, elided or omitted, according to the needs of the topic and the purpose of the text.

The subject proposed.

I. Whatever is needed to explain the subject

Repetition

Obverse

Definition

II. Whatever is needed to establish the subject

Exemplification or detail

Illustration

Proof

III. Whatever is needed to apply the subject

Result or consequence

Enforcement

Summary or recapitulation

Canby (1909) repeated Genung's formula, but modified it a little by sub-classifying definition as limitation, restriction or enlargement, and by stating that the subject is usually proposed in the form of a topic sentence. The only other of our writers to present such a formula is Thompson (1957), but his, which follows, is much less detailed - so much so that for teaching purposes it is virtually useless.

1. Topic sentence (central idea)
2. Series of additional sentences supplying details
3. Final sentence, summarising, applying, and/or generalising

Kane and Peters (1966) list three basic techniques, giving examples and hints for implementing them :

1. Amplification (Illustration, Analysis, Restatement)
2. Comparison
3. Logical techniques (Cause and Effect, Definition, Qualification, Negative Development, etc.)

This approach differs from that of Genung or of Thompson, in that the writer may use only one of these methods, and if he uses more is free to order and arrange them in any way he pleases.

Although all these writers accept the prevailing paradigm of sentence grammar, they believe at the same time that some kind of order exists at higher levels, and that a rhetorical patterning is present, and can, up to a point, be described. They go to some lengths to list communicative acts, but their failure to include formal descriptions of the linguistic manifestations of these acts renders their textbooks relatively unhelpful as far as the student is concerned. Some writers do try to solve this problem by presenting copious examples, but these examples frequently fail to illustrate the writer's theories. Genung, for example, uses illustrations which deviate considerably from his formula, and if he had carefully studied much authentic text he would have seen that most paragraphs do not follow such a rigid pattern. Nevertheless, Genung should be given credit for at least attempting to see rules behind paragraph structure, and it is interesting that eighty years later a very similar, but much more simplistic, set of formulae appeared as a new theory of the paragraph, within Pike's Tagmemic Model. (See Chapter 3).

Jones and Faulkner (1961) are the only authors here to attempt a full-scale model of the development of the text from the sentence, via the paragraph, to the complete theme. This part of their book, however, is outside the mainstream of traditional rhetoric, having been strongly influenced by structural linguistics, and so it will be dealt with in the next chapter, when we consider some developments of the sixties.

3.6 Sentence length Almost every writer has some comment to make about the length of sentence within the paragraph. A major problem here as far as the learner is concerned, is that the advantages and disadvantages of long and short sentences are usually discussed with no indication given as to the length of a "long" or a "short" sentence. It is implied that every sentence must be "long" or "short", whereas the majority of sentences in actual text probably come at some point in between.

The commonest advice in these textbooks from Bain onwards is to vary sentence length so as to avoid monotony. Baldwin (1902) diverges from the common consensus when he recommends the use of long sentences rather than short because of their "easier flow" (1902 : 2), and Bain himself gives as one of the advantages of long sentences the fact that "they serve to group related facts", and they "permit the expansion of a thought". (1866 : 84-5). On the other hand, short sentences are seen as being "simple and direct" (Bain 1866 : 84) "impart vigour", and "give point and crispness to a thought" (Genung and Hansen 1915 : 152).

Although variety is almost unanimously recommended, few of these writers indicate where or when the short sentence, or indeed the long sentence, should be used. Again Genung, one of the earliest writers, gives the most detailed and specific information. He explicitly denies the possibility of prescribing any definite length for a sentence, saying that this should be left in all cases to the writer's "discretion and taste", but he does make three very important points, which will be quoted in detail.

> 1. The short sentence, being easier to understand and remember, is especially adapted to passages where important points have to be made, passages of definition or discrimination, or on which much of the thought hinges. The fundamental propositions which constitute the central nucleus of a course of thought, and passages of summary, are generally expressed in short sentences. The short sentence may often be used to advantage, also, for purposes of emphasis, the successive condensed assertions being like so many hammer strokes ...
> 2. The long sentence affords room to amplify the sense, by considerations ancillary to the main idea ; it is therefore serviceable for introducing details filling out a previously suggested thought ...
> 3. Between long sentences of detailed thought it is often necessary to insert short transitional sentences suggesting the thought in sententious form. There is perhaps no more fruitful cause of "hard reading" than the neglect to supply such compendious means of transition and connection.
>
> (Genung 1915 : 185-187)

The first part seems to suggest that the topic sentence, and maybe the last sentence of a paragraph, should on occasion be short. If the topic sentence presents the main idea of the paragraph, then the sentence or sentences immediately following it, adding supporting facts, should be long. Most important is the third point, that short sentences

may be used to indicate a transition, perhaps within a paragraph, and that this will help to ease the reading burden.

It is surprising that an aspect of writing which was clearly seen by Genung as being of great importance was virtually neglected by most later writers, although Smart did suggest that "a single thought may be made to stand out prominently by placing it in a single short sentence" (Smart 1922 : 137). Indeed, among the texts now being examined, it was not until over seventy years later that Kane and Peters (1966) wrote in similar terms to Genung. They divide sentences into segregating sentences, which contain only one "idea", and aggregating sentences containing several "ideas". No attempt is made to define an "idea", but we can perhaps guess that a segregating sentence is what the traditional grammarians would call a "simple sentence", ie a sentence consisting of one main clause only, while an aggregating sentence is either a "complex sentence", containing at least one main clause and one or more subordinate clauses, or a "compound sentence", containing more than one main clause. Kane and Peters point out that the segregating style is common in children's stories and some kinds of narration and description, but not in expository writing. The isolated segregating sentence, however, should be used, particularly initially, and where there is a transition of thought.

> The isolated segregating sentence, then, if not the segregating style, is valuable to the expository writer when forcefulness and lucidity are cardinal considerations. It is especially useful in topic sentences and in transitions, in emphatic restatements within the paragraph, and less commonly in closing sentences. It is exceedingly difficult to write exposition effectively in segregating sentences for any length of time. But it is equally true that it is difficult to write well for any length of time without using a segregating sentence now and then. If for no other reason, short sentences should be called upon to break up the monotony of longer constructions. The able writer, however, will do more than this. He will make the segregating sentence do double duty, using it not only to vary the rhythm and movement of his style, but at the same time to emphasize an important point.
>
> (Kane & Peters 1966 : 256)

3.7 Paragraph length Little more specific guidance is given concerning paragraph length than concerning sentence length. All the writers who deal with this matter agree that readability and appearance are just as important as the topic under discussion. Paragraphs should therefore not be too long. Hulbert and Hulbert suggest :

> In developing a topic into a paragraph the student should watch its length, and if he finds his discussion becoming too long for one paragraph he should divide it into two or more paragraphs.
>
> (Hulbert & Hulbert 1929 : 34)

Kierzek (1939) similarly suggests, somewhat vaguely, that although very short paragraphs should be avoided, if paragraphs become too long they should be divided up. Genung points out also that too much solid print with few paragraph breaks will discourage the reader, whereas "an open and easy-looking page" will attract him to read on to the end.

No writer is very anxious to prescribe the optimal length for an ideal paragraph, seeing this as a matter of the writer's individual choice, and depending on such factors as audience and subject matter. According to Brooks and Warren :

> .... there is no precise formula by which the length or structure of a paragraph may be determined .... The writer must use his best judgment : he must use his common-sense and his taste. Unless he is very sure of his ground, he will tend to employ paragraphs of medium length.
>
> (Brooks & Warren 1950 : 299)

A few writers do stipulate paragraph length more exactly. Genung, for example, suggests that a paragraph should never be longer than one page, but he does not consider the obvious fact that pages are of different sizes, and that therefore what he is really implying is that paragraphs can vary in size according to the size of the page. Hulbert and Hulbert (1929) take this a stage further. They advocate paragraphs of between a half and one-and-a-half pages in length, but also point out that paragraphs written in longhand are usually shorter than those in typescript, and these are in their turn shorter than printed paragraphs. Kane and Peters (1966) carry this idea to its ultimate conclusion by recommending that a typewritten paragraph should be between a half and a third of a page in length, while a handwritten paragraph can be between one and two pages.

Although most of these writers recommend variety in sentence length, not one of them is in favour of variety in

paragraph length. Shaw (1946), however, does point out that very short paragraphs are sometimes acceptable for emphasis, or for a particular stylistic effect.

4. Evaluation

The approaches of traditional rhetoric tend nowadays, at least in Britain, to be ignored. The above survey should show, however, that, as with traditional grammar, by indiscriminately discarding everything much of value has been lost. No other approach has yet produced a greater wealth of ideas.

Perhaps one of the drawbacks of traditional rhetoric as a teaching tool is this very wealth of ideas. The majority of the books surveyed present what seems endless prescriptions and model passages. The vital missing element is pedagogical, some method concerned with more than simple assimilation, to bridge the gap between learning about rules and learning how to apply them. It is true that most of these books contain exercises, often not very different from those considered "modern" today, but, in every case except perhaps the most recent, theory bulks larger than practice.

The traditional rhetorician's subjective approach, however, is the main reason for his loss of favour. His precepts are based not on analysis and experiment but on intuition, and on the repetition decade after decade of the ideas of previous writers. His rules often seem to apply to some kind of pseudo-text rather than to genuine writing. The language learner has to assimilate some rules, but we

should at least make sure that these rules are applicable to normal, authentic text.

Nevertheless, on examining the ideas of more modern schools of thought, it becomes clear that the more scientific approach has produced very little that is genuinely new. Rhetoric has had a start of centuries, and in that time has stated, suggested or implied much of what is claimed as "new" by those concerned today with the study of text structure and the production of teaching materials for reading and writing skills. One important insight, however, which the traditional rhetoricians seem not to have stumbled on is the hierarchical structuring of expository text, if not of all text. This feature will be one of those considered in the next chapter, where more modern developments from traditional rhetoric will be considered.

CHAPTER 3

THE CONTRIBUTION OF AMERICAN TEACHERS OF COLLEGE COMPOSITION IN THE SIXTIES

1. Introduction

The writers to be considered in this chapter are all practical teachers of rhetoric working in the United States in the 1960s and early 1970s. Their ideas descend directly from those of the traditional rhetoricians, but diverge from them on account of dissatisfaction with the effectiveness of unformalised traditional prescriptive rhetoric as a teaching tool. Although most are not trained linguists, all would acknowledge a debt to the linguistic theories, both structuralist and transformational-generative, prevalent in the 50s and 60s.

The link between all these writers is their attempt to produce models of the surface structure of expository prose in English. At this time most American linguists, with the exception of the Tagmemicists, were still exclusively sentence orientated, so it was inevitably left to non-linguists to enter the arena of discourse. Their purpose was practical rather than academic. They believed that if their students could be given an efficient method for the analysis of "good" prose, they would learn as a result to produce better prose on their own account.

The earlier writers in this section approached expository prose from the point of view of the paragraph,

but as they were forced to a greater and greater extent to consider text as an organic whole, they moved away from the analysis of isolated paragraphs to the analysis of total structure.

2. Francis Christensen

During the 1960s considerable interest in the teaching of writing skills was aroused in the United States, as can be seen from the numerous articles and discussions appearing in the journal "College Composition and Communication". This interest and concern arose from the failure of many students in American universities to benefit from compulsory freshman writing courses, and subsequent attempts by their teachers to find new and more successful approaches. It was further encouraged by the radical changes taking place in the academic discipline of linguistics, and the advocacy of a "new rhetoric" by such scholars as Kenneth Burke and I.A. Richards. The earliest and most influential teacher and writer to become well-known in this field was Francis Christensen of the University of Southern California.

It should be emphasised that Christensen and his followers (eg Karrfalt, Pitkins, Miles and Larson) were trained primarily in literature rather than in linguistics or psychology, though they claim to have been influenced by Chomsky's theories of transformational-generative grammar. Their main concern was the teaching of literary and expository writing to American undergraduates, and their work was not experimentally based. Nevertheless, Christensen and his

followers and critics have produced insights into the structuring of prose which are of considerable value for both the teacher and researcher, and a critical resumé of their work will therefore be given.

2.1 Christensen versus Hunt : the cumulative sentence.

Christensen himself was led to his theory of the paragraph via his consideration of the structure of the sentence. By 1963, when Christensen wrote his first important article, Kellogg Hunt had already begun his own analysis of school-children's essays, to find out what constituted "syntactic maturity" in writing. Hunt's research suggested that as students become older and more mature they tend to write longer clauses with a greater use of subordination, at least up to the middle grades. He stated that at a later stage clauses become reduced to single words and phrases, for, ".... as the mind matures, it organises information more intricately, and so can produce and receive more intricately organised sentences". (Hunt,1970). Christensen strongly attacked what he considered to be Hunt's confusion between complexity of structure and complexity of thought. According to his view :

> A mature style must say much in little, agreed, but a mature style must be easy to decode. The long clause is not the mark of a mature style, but of an inept style - the easy writing that's curst hard reading. The real problem in writing is to reconcile these two seeming opposites - to pack much into little, but to pack it so that it is easily unpacked.
>
> (Christensen 1968 : 576)

Christensen's answer to Hunt was the advocacy of the cumulative sentence.

It should be noted that Hunt, the psychologist, was judging maturity by the progression that his experiments suggested was normally made in a child's writing as he grew older. Christensen, the literary scholar, was equating "mature" writing with "good" writing. He was looking at exceptional literary prose, such as that of Hemingway, so he naturally saw a different type of prose from Hunt. The pedagogical application of Hunt's theory (eg as attempted by Mellon (1967))would involve speeding up a child's natural development in writing prose, while the application of Christensen's theory would involve teaching children, or preferably older students, to produce a new type of sentence on the model of selected "good" modern writers.

Christensen spoke about a "generative" rhetoric of the sentence, and later of the paragraph, but one should not be misled into supposing that this was in any way connected with transformational grammar. It is true that Christensen and his followers were great admirers of Noam Chomsky, but "generative" (incidentally a very fashionable word in the sixties) here means simply that the rhetoric they suggested would generate ideas, and not merely combine the ideas of simple sentences.

The cumulative sentence was the result of the application of the following four principles. These are particularly important, because two years later Christensen went on to apply the same principles to the paragraph, which he considered as a "macro-sentence".

1. Addition Christensen quoted John Erskine (1946): "When you write you make a point, not by subtracting as though you sharpened a pencil, but by adding".
2. Movement or direction of modification. If a modifier is added before the head, the direction of modification is forwards, and if it is added after the head the direction is backwards. Modification is best made after the clause rather than within the clause, as the resultant "cumulative sentence" will then grow as the mind thinks, and be more dynamic. Movement within the sentence will be constantly ebbing and flowing - forward to a new point, backward to modify the statement of the main clause, and then forward again.
3. Levels of generality The main clause is usually the most general, with succeeding clauses particularising the main clause. Christensen gives an example of this : "He has just bought a new car, a 1963 Ford, a Galaxie, a fastback hardtop with four-on-the-floor shift" (Christensen 1963 : 157).
4. Texture This should be dense rather than thin, with considerable added modification.

Christensen saw the cumulative sentence as an alternative to the type of sentence that Hunt seemed to be advocating, the sentence containing long noun phrases. He

gives examples of both his own type and Hunt's (Christensen, 1968 : 575) :

1. Sentence with noun phrases

   The curriculum is at best, however, a design to be interpreted by teachers with varying degrees of ability and insight for children with differing equipment in intelligence and language background.

2. Cumulative sentence

   The curriculum is at best, however, a design to be interpreted by teachers for children - by teachers with varying degrees of ability and insight, for children with differing equipment in intelligence and language background.

2.2 Application of the cumulative sentence to the paragraph

In 1965 Christensen wrote an article entitled "A Generative Rhetoric of the Paragraph". In this he aimed to show that his four principles of sentence development also apply to the paragraph, with the topic sentence parallel to the base clause, and the supporting sentences equivalent to added modifiers.

In trying to work out a model for the paragraph, Christensen could not avoid a problem central to all study of text organisation above the sentence - the fact that as one ascends the hierarchy dependence on rules seems to lessen and exceptions increase, until one begins to doubt whether a rule-based model is really valid. Any attempt, therefore, to fit a model to all existing paragraphs must

inevitably fail, and some of the criticism levied against Christensen (eg Rodgers 1966), on the grounds that he did not do just this, was unreasonable. To allow for exceptions he knew existed Christensen suggested five rules which always apply, followed by four "rules" to cover less usual types of paragraph.

2.2.1 Five rules of paragraph development The first five rules were as follows (Christensen, 1965) :

1. The paragraph may be defined as a sequence of structurally related sentences.
2. The top sentence of the sequence is the topic sentence.
3. The topic sentence is nearly always the first sentence of the sequence.
4. Simple sequences are of two sorts : coordinate and subordinate.
5. The two sorts of sequence combine to produce the commonest sort : the mixed sequence.

By means of a very simple analogy between the sentence and the paragraph (which of itself was not new, as we have seen), Christensen explained at one blow two important features of the paragraph which had been recognised but unformalised for nearly a hundred years. First, the unity of the paragraph, which, like the unity of the sentence, he claimed depended on coordination and subordination. Second, the topic sentence, previously recognised because of its meaning, now defined as the sentence upon which the other sentences depend, in the same way that coordinate and subordinate clauses and added modifiers depend on the main clause of a

sentence. Implicit in Christensen's model, though nowhere actually stated, is the belief that discourse is hierarchical in structure.

Christensen gives a number of examples of paragraphs analysed according to this system. He uses a very simple graphical method : the topic, or first level sentence, is always numbered as 1, with succeeding levels numbered accordingly. Every coordinate sentence is placed immediately beneath the sentence to which it is related, while subordinate sentences are placed beneath, but slightly to the right. One of his examples can be seen in Figure 4.

A major weakness is revealed, however, when one attempts to put Christensen's model into practice. Nowhere does he state in exact terms how one should decide whether a sentence is coordinate or subordinate. He suggests a provisional assumption that the first sentence is the topic sentence, and then a search for "likeness", indicating coordination. He never makes clear, however, what this "likeness" consists of, mentioning grammatical features, eg identity of structure at the beginning of a sentence, and also meaning and function, eg when he suggests that coordinate sentences are added to enumerate or emphasise. He points out that there are certain clues or signals to coordination which the analyser should look out for, including such paired words and phrases as "We can hardly say / Nor can we argue", or "used to be / now". Subordination he hardly deals with, assigning it to a "ragbag" for whatever is left over when the topic and coordinate sentences have been dealt with.

FIGURE 4

CHRISTENSEN'S ANALYSIS (1)

1) 1. The purpose of science is to describe the world in an orderly scheme or language which will help us to look ahead.

2) 2. We want to forecast what we can of the future behaviour of the world ; particularly we want to forecast how it would behave under several alternative actions of our own between which we are usually trying to choose.

3) 3. This is a very limited purpose.

4) 4. It has nothing whatever to do with bold generalisations about the universal workings of cause and effect.

5) 4. It has nothing to do with cause and effect at all, or with any other special mechanism.

6) 4. Nothing in this purpose, which is to order the world as an aid to decision and action, implies that the order must be of one kind rather than another.

7) 5. The order is what we find to work, conveniently and instructively.

8) 5. It is not something we stipulate ; it is not something we can dogmatise about.

9) 5. It is what we find ; it is what we find useful.

From J. Bronowsky The Common Sense of Science

2.2.2 Four modifying rules of paragraph development. In order to cover exceptions, Christensen suggested the following modifying rules :

1. Some paragraphs have no topic sentence.
2. Some paragraphs have sentences at the beginning or at the end that do not belong to the sequence.
3. Some paragraphing is illogical.
4. Punctuation should be by the paragraph and not by the sentence.

The first point only makes sense when we consider Christensen's belief that a new paragraph begins when a succeeding sentence is neither subordinate nor coordinate with that preceding it. Thus, in Figure 4, if the next sentence were at level 5 or 6 it would belong to the same paragraph. If, however, it were to be at any higher level, it could, or should, start a new paragraph. Presumably the higher the level the more likely it would be for a new paragraph to begin.

Some paragraphs, however, do not begin at level 1, and an example of Christensen's can be seen in Figure 5. Here a new paragraph begins at level 3, levels 1 and 2 forming the preceding paragraph. The two paragraphs taken together follow Becker's pattern of Topic - Restriction - Illustration (See Section 7.1), and could have formed one unified paragraph. The author, however, preferred to split it into two, the first part being the introduction (and thus the topic for both paragraphs) and the second part the illustration. She may have considered the passage too long for one paragraph alone, or have wished to highlight the examples by paragraphing them separately.

FIGURE 5

CHRISTENSEN'S ANALYSIS (2)

1) 1. The mystic artist always sees patterns.

2) 2. The symbol, never quite real, tends to be expressed less and less realistically, and as the reality becomes abstracted the pattern comes forward.

3) ¶ 3. The wings on Blake's angels do not look like real wings, nor are they there because wings belong to angels.

4) 4. They have been flattened, stylized, to provide a curving, pointed frame, the setting required by the pattern of the composition.

5) 3. In Hindu art and its branches, stylization reached its height.

6) 4. Human figures are stylized far beyond the point of becoming a type ; they too are made into patterns, schematic designs of the human body, an abstraction of humanity.

7) 3. In the case of an Eastern rug, all desire to express any semblance of reality has gone.

8) 4. Such a work of art is a pure decoration.

9) 5. It is the expression of the artist's final withdrawal from the visible world, essentially his denial of the intellect.

From Edith Hamilton The Greek Way

Notice that the separating off of the highest levels is common in introductions, in which case the whole of the rest of the passage, however many paragraphs it may consist of, is at a lower level. If Christensen had considered whole discourses, eg essays, chapters, or complete books, rather than paragraphs in isolation, he would have realized that very few paragraphs truly begin at level 1. The picture is far more complicated, and Christensen realizes this, but does not pursue his investigation any deeper.

A further difficulty Christensen faces is the result of writing skills being specialized, and learnt relatively late in life (compared with oral skills) by an élite minority. It is possible to assume that all educated speakers of English as a mother tongue will write sentences that are rule-governed, and can be considered as "grammatically correct". It is not possible to assume this about rhetorical skills above sentence level. Many examples of illogical and obviously incorrect paragraphing can be found, even among reputable and well-known writers. Christensen notes this in Rule 3 above, and gives an actual example from C. Day Lewis.

2.3 Evaluation This model has inspired practical teachers of rhetoric in America for a decade or more, and has been used by numerous researchers, but in spite of this it remains a very inadequate method when actually applied to expository prose, apart from the carefully chosen examples Christensen uses. To illustrate some of the difficulties, the first part of one of the three texts to be used in later experiments, a passage

from L. Dudley Stamp's "Our Underdeveloped World", will be analysed according to Christensen's model. The full text appears in Appendix 4 and the analysis in Figure 6. This passage, which from now on will be referred to as the soil erosion passage, forms a satisfactory whole, even though it is only a small part of a complete book. It displays both coordination and subordination, but very few explicit signals are used for either - analysis is dependent almost totally on meaning. Criticism of Christensen's model as applied to this particular passage can be summarised as follows:

(i) The analysis is subjective, and areas of doubt can be found where alternative explanations are possible. For example, a geographer who examined the passage doubted whether the last of the level 3 coordinate sentences (sentence 7) could be grouped with the others, on the grounds that the wind is not an erosive force but an agent of transport. Christensen suggests that the analyst should go through the passage and look for signs of likeness, but in expository texts intended for mother tongue speakers explicit signalling is usually limited. In this passage the interpretation of sentences 3 to 7 as a coordinate block depends on the reader's knowledge concerning agents of erosion. It is true that a certain amount of structural parallelism is evident in sentences 3 and 4, but this breaks down completely in sentence 5. Indeed, when an attempt was made to get highly educated subjects to divide the passage up into groups of sentences below paragraph level, only a bare majority of them chose to

FIGURE 6

CHRISTENSEN'S ANALYSIS : THE SOIL EROSION PASSAGE

1) 1. Every geologist is familiar with the erosion cycle.
2) 2. No sooner has an area been raised above sea-level than it becomes subject to the erosive forces of nature.
3) 3. The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets and then into rivers and out to sea.
4) 3. The frost freezes the rain water in cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust.
5) 3. Blocks of rock dislodged at high levels are brought down by the force of gravity.
6) 3. Alternate heating and cooling of bare rock surfaces causes their disintegration.
7) 3. In the arid regions of the world the wind is a powerful force in removing material from one area to another.
8) 4. All this is natural.
9) 2. But Nature has also provided certain defensive forces.
10) 3. Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks.
11) 4. Slowly but surely different types of soil, with differing profiles, evolve, the main types depending primarily on the climate.
12) 3. The protective soil covering, once it is formed, is held together by the vegetation.
13) 4. Grass and herbaceous plants, with long, branching, tenuous roots, hold firmly together the surface particles.
14) 3. The same is true with the forest cover.

15) 4. The heaviest tropical downpours, beating on the leaves of the giant trees, reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil.

16a) 3. The soil, thus protected by grass, herbs and trees, furnishes a quiet habitat for a myriad varied organisms:

16b) 4. earthworms that importantly modify the soil, bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants.

17a) 3. Chemical action is constantly taking place :

17b) 4. soil acids attack mineral particles, and salts in solution move from one layer in the soil to another.

18) 2. We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions and parent rock substances develops.

19a) 3. It is a long process, and in many areas there has not been sufficient time in the geological sense for the completion of the process :

19b) 4. the soils are immature.

treat this group as a single unit. As Larson (1967) points out, it is necessary to make the relationships between the sentences much clearer, possibly using a limited number of labels, if we want a more effective analysis than Christensen's.

(ii) Christensen is looking for consecutive sentence-by-sentence relationships, and ignoring groups of sentences which hang together as units between sentence and paragraph level. For example, sentence 8 is not subordinate to sentence 7 alone. It could be considered either as subordinate to the whole block (sentences 3 to 7) or to sentence 2, in which case it would be at the same level as the block, which is clearly wrong. A similar problem arises with sentence 16a, a problem less easily resolved. Here, the sentence is clearly not dependent on the previous sentence, but on two closely related groups of sentences, 12 to 13 and 14 to 15. The main sentences of each of these have been placed at level 3, so sentence 16a could therefore be at level 4. If so, however, it would then appear to be coordinate with sentence 15, which is clearly false. Level 3 has in fact been chosen because a new point is being made about the soil, but no judgment could be entirely satisfactory.

(iii) The summary sentence presents a similar problem. The summary is in one sense subordinate to the whole passage, so could be placed at level 5, but this would suggest erroneously that it is subordinate to sentence 17. It

has therefore been placed at level 2, to indicate that it is subordinate to the first sentence, as is the rest of the passage.

(iv) Christensen works with complete sentences, however long they are and however many main clauses they may contain. This inevitably results in a concealment of some of a passage's structure, and so in this text main clauses separated by colons and semi-colons have been treated as separate sentences (sentences 16, 17 and 19). It might produce better results to analyse a text in terms of T-units rather than sentences (Hunt 1970b)

3. David Karrfalt

Some of the above criticisms have been made, and revisions suggested, by colleagues and students of Christensen, among whom can be numbered David Karrfalt.

3.1 Revision of Christensen's model. Karrfalt (1968), was at first particularly concerned with the problem of the summary sentence, a sentence which is usually at a higher level of generality than not only the preceding sentence, but sometimes most of the paragraph. When he tried to apply the model to a number of different paragraphs in texts, he found that often a sentence that seemed to break away from the coordinate-subordinate sequence was also, like the summary sentence, at a higher level of generality. An example can be seen in Figure 7.

FIGURE 7

KARRFALT'S ANALYSIS

1) 1. There is an amazement proper to the experience of all great art, but the special amazement which "War and Peace" revives in me while I am reading it is like that of a child.

2) 2. The child does not expect the unexpected ; that would already be a preparation against it.

3) 2. He does not for an instant doubt that a certain event had to happen ; such doubt obscures.

4) 3. He may even have been told beforehand that it was going to happen ; such foreknowledge is as little a part of him as is a label in his cap.

5) 3. He is able to look at the thing itself.

6) 3. The event reaches him radiant with magical causes, but not yet trapped in sufficient cause.

7) /1. Tolstoy does not, as many do, achieve this freshness by transforming the reader into a never-never land.

8) /1. On the contrary, his fictional mode is realistic : the people in his novel appear and behave like possible people in the world we daily live in.

9) /2. His achievement is the greater because he uses the mode of realism, for realism offers a threat to which other literary modes are not subject, the encroachment of mediocrity.

From George Elliot A Piece of Lettuce

Karrfalt faces the same problem here that we have just been considering. Sentence 7 is clearly not subordinate to sentence 6. It is obviously structurally related to the first sentence, but if placed as subordinate to it would appear coordinate with sentences 2 and 3, which it is not. It is in fact related in some way to all the preceding sentences in the paragraph.

Karrfalt solves the problem in the following way. He feels that the concept of levels of generality is basic to Christensen's model, but that one important dimension has been left out. A sentence may be at the same level of abstraction as its predecessor, in which case the relationship is one of coordination. It may be at a lower level, when the relationship is one of subordination. It may, however, be at a higher level, a relationship which Karrfalt labelled as completion, and in this case it may be at a higher level not merely to the preceding sentence but also to all the foregoing sentences taken as a single unit. Sentences with this relationship are represented by Karrfalt as in Figure 7.

3.2 Evaluation. While performing a useful purpose in highlighting a problem, this solution does not succeed in resolving the confusion. It would seem more appropriate for sentence 7 to be made subordinate to sentence 1, as "this freshness" refers to the childlike amazement described in the first sentence. Christensen gives us no reason for assuming that all sentences at level 2 are coordinate : nevertheless, one could agree with Karrfalt up to a point, in that it is aesthetically unsatisfying if this is not so. A better

solution is that mentioned above - to devise some form of labelling, so that it is obvious that, although at the same level, sentences 2 and 3 are coordinate and sentence 7 is not.

Whichever way Figure 7 is analysed, however, it can be said that the possibility of a new paragraph at sentence 7 is present, as this sentence is neither coordinate nor subordinate with what precedes it. A conclusion can probably always be separated off as a separate paragraph, the restraint being the optimum paragraph length for the particular type of format. (See Chapter 8).

4. Willis Pitkins

Another student of Christensen, Willis Pitkins, was the first member of this group to make explicit the unspoken assumption that discourse is hierarchical in structure. In an important 1968 paper he wrote :

> ... the structure of written discourse - like the structure of the complex word, the phrase, the clause - is hierarchical, units embedded within or added to larger units embedded within or added to still larger units ; and at any level of the continuum the units are to be discovered not by their shape on a page, not by how long they are or how they are punctuated, but by what function they are serving in the discourse. And as there are no structural gaps in what we recognise as well-formed words or sentences, so there are no structural gaps in what we recognise as well-formed discourse.
>
> (Pitkins 1969 : 141)

Here Pitkins makes it clear that paragraphs cannot be considered in isolation : every paragraph is a part of a larger structure, indeed, perhaps a very small part if the complete discourse happens to be a full length book.

By looking for the first time at the complete discourse, however, Pitkins is caught in a dilemma which he himself does not fully recognise. As a practicing teacher of composition, Pitkins needs to show his students how to produce work written in logical, well-developed paragraphs,ie to teach them when to indent and what to put between the indentations. When he looks at the structure of authentic discourse, however, rather than at isolated paragraphs, he sees units of all sizes embedded in each other like Chinese boxes. One solution might be to abandon the paragraph as a conceptual reality, retaining it only as an orthographic device to help clarify the text. Pitkins does not go as far as this, but instead he introduces a new unit, parallel to the sentence and the paragraph, which he calls the discourse bloc.

4.1 Pitkin's model Pitkins criticises Christensen's model on two main counts. First, it is sentence-orientated. Pitkins sees this as the result of neglecting function : some sentences have only one function in the discourse, and in this case they may be considered as wholes, but where a sentence has two or more functions then it should be subdivided accordingly.

More seriously, he finds the basis upon which sentences are labelled as coordinate or subordinate to be inadequate. Not all relationships fit into this pattern, eg positive/ negative, or question/answer. He proposes instead the replacement of coordination by horizontal relation and of subordination by vertical relation. The horizontal relation

would consist of coordination (as proposed by Christensen) and complementation (to cover the above named relationships) The vertical relation would consist of subordination and superordination (ie Karrfalt's 'completion').

Pitkins states his hypothesis as follows :

> Connected discourse is a hierarchy of discourse blocs. Discourse blocs, whatever their form, are function units. At any given level of the hierarchy there will be only two blocs, unless the level represents a coordinate series, which, theoretically at least, can have any plural number of members. And at any given level the blocs will be related according to one of four possible broad relations - two vertical (superordination and subordination), and two horizontal (coordination and complementation).
>
> (Pitkins 1969 : 142)

The graphical representation of this theory is both more sophisticated and more informative than Christensen's. Pitkins himself uses the same passage as Christensen for the purpose of illustration. (See Figure 8). In this fairly simple example it can be seen that the divisions are binary, except in the case of coordination. The type of relationship is symbolised by the four prosidic feet, ie subordination (⏜ ⏝), superordination (⏝ ⏜), coordination (⏝ ⏝), and complementation (⏜ ⏜), and each relationship is specifically named at the top of the relevant column. The vertical lines stop short at the horizontal in order to leave room for the placing of any explicit "bloc signals" over them.

Such a diagram can only illustrate a very small part of a complete text, but it makes clear the close inter-relationships present. In this example we may consider the complete paragraph as a discourse bloc. Within it we can

FIGURE 8

PITKINS' ANALYSIS

assertion of purpose			categorisation of purpose						
assertion		PARTICULARLY particular	categorisation	support					
assertion	repetition			negative dimension			positive dimension		
				assertion		reassertion	assertion		reassertion
				assertion	generalisation		assertion	neg of contrary	
SENTENCE 1	2a	2b	3	4	5	6	7	8	9

identify two smaller blocs, sentences 1 and 2, and sentences 3 to 9. If we then consider the second bloc, it can be sub-divided further, into negative and positive dimensions (Sentences 4 to 6 and 7 to 9). In turn, each dimension can be divided into still smaller blocs, eg the positive dimension into assertion (sentences 7 and 8) and reassertion (sentence 9). There is no need to stop at sentence level. Pitkins also shows how a similar analysis can be made at clause level.

4.2 Evaluation It is possible to analyse the soil erosion passage according to Pitkin's model (Figure 9) and compare it with the Christensen analysis (Figure 6).

The functions are provided with labels used by Pitkins himself in his examples, and the fact that this did not prove difficult suggests that a drawing up of a limited workable list is feasible. This system gives an immediate picture of the structure of the passage as Christensen's analysis does not. It is possible to take any sentence and work backwards, level by level, to find out exactly how it fits into the whole. The addition of the vertical relation of superordination resolves the problem of the summary (sentences 18 and 19), and similarly the relationship of complementation between sentences 16 to 17 and sentences 10 to 15 is made clear. A minor criticism might be that the diagram does not show the hierarchical organisation of the text as well as it might, in that whatever the levels of the binary divisions are in relation to each other, they are always placed graphically at the same level, ie side by side. This is perhaps the sole

FIGURE 9

T.S. ASSERT	SUPPORT																											
	NEGATIVE DIMENSION								BRIDGE	POSITIVE DIMENSION																		
	T.S. ASSERT	SUPPORT								T.S. ASSERT	SUPPORT																	
		EXAMPLE		EXAMPLE		EX	EX	EX			PARTICULARS														GENERALISATION			
		PROC 1	PROC 2	PROC 1	PROC 2						PROCESS							RESULT							ASS	CATEGORISATION		
											PROCESS 1		PROCESS 2					ASSERTION			ASSERTION				ASSERT		RES
											ASS	SUPP	ASSERT		COMPARISON			ASS	PART		ASS	PART			ASS	RES	
													ASS	PART	ASS	SUPPORT			EX	EX		EX	EX				
																PROC 1	PROC 2	PROC 3										
1	2	3a	3b	4a	4b	5	6	7	8	9	10	11	12	13	14	15a	15b	15c	16a	16b	16c	17a	17b	17c	18	19a	19b	19c

PITKIN'S ANALYSIS :

THE SOIL EROSION PASSAGE

disadvantage that this system has in relation to Christensen's model.

Pitkins has not been able to overcome the perhaps inevitable element of subjectivity involved in such an analysis. He has reduced the subjective element, however, by forcing the analyst to apply labels to his choices, and thus think more clearly about the functions performed by the blocs. Unfortunately he makes no attempt to list the types, and to limit them to any particular number, his only comment on this being that the final number is probably "workable".

5. Robert Kaplan

Robert Kaplan, also from the University of Southern California, was one of the first teachers of English as a foreign language to turn his attention to the special problems of writing expository prose. His main research area is that of contrastive rhetoric, but he became interested in the work of Christensen and Pitkins, and explored the possibility of using the rhetorical analysis of texts as a teaching instrument for foreign learners, in the same way that it was being used for mother-tongue speakers.

5.1 Kaplan's model Kaplan (1970) was particularly impressed by Pitkins' theory of discourse blocs. His interpretation of the term, however, is not exactly what Pitkins seemed to mean by it in his 1968 paper. Pitkins sees the discourse bloc as a unit parallel to the sentence and the paragraph, whereas Kaplan sees it as a unit contained within the same system as the paragraph, though not necessarily in any way related to

the sentence. Pitkins sees discourse blocs as fitting inside one another like Chinese boxes. He does not suggest any upward limit on size, the complete discourse presumably forming the first bloc at the highest level, while in one example he sub-divides a section of the text down to single words. Kaplan's interpretation of Pitkins is obscure, but he seems to be suggesting that a discourse bloc is in some way equivalent to a paragraph (though perhaps a conceptual rather than an orthographic paragraph (see Lackstrom Selinker and Trimble 1970, 1973)).

Kaplan therefore suggests that paragraph analysis would be easier if a unit below the level of the bloc were introduced, a unit which he terms the discourse unit. He defines the two terms as follows : (Kaplan 1970 : 51)

> The discourse bloc : A unit "in which every item is related to every other item by either coordination, subordination, or superordination".
>
> Discourse units : "Those units within a discourse bloc which are related to each other by either coordination, subordination, or superordination".

Kaplan does not seem to recognise that every discourse bloc below the level of the complete text is, according to this definition, also a discourse unit, while every discourse unit except the smallest, by virtue of containing coordinated, subordinated or superordinated units within it, is also a discourse bloc. He does not present examples of more than one paragraph in length, so that the impression is given that the paragraph is a complete and self-sufficient unit, which can

FIGURE 10

KAPLAN'S ANALYSIS

DISCOURSE BLOC

I Discourse Unit 1 (prefatory statement /topic sentence)

1) The purpose of science is to describe the world in an orderly scheme of language which will help us look ahead.

II Discourse Unit 2 (subordinate modifier of Unit 1)

2a) Head : We want to forecast what we can of the future behaviour of the world ;
(subordination keyed on semantic structure "look ahead, forecast)

2b) A. Coordinate modifier (keyed on repetition/ reassertion of structure "we want to forecast") particularly we want to forecast how it would behave under several alternative actions of our own between which we are usually trying to choose.

III Discourse Unit 3 (subordinate modifier of Unit 2)

3) Head : This is a very limited purpose.
(subordination keyed on pronoun reference "this")

4) A. Subordinate modifier of head (keyed on pronoun reference "it")
It has nothing whatever to do with bold generalisations about the universal workings of cause and effect.

5) B. Coordinate modifier with A (keyed on repetition of same pronoun reference "it" and repetition of structure "it has nothing/ whatever/to do with" : It has nothing whatever to do with cause and effect at all, or with any other special mechanism.

6) C. Coordinate modifier with A (keyed on pronoun reference "this") :
Nothing in this purpose . . . . implies that the order must be of one kind rather than another.

6(cont) 1. Subordinate modifier of C (keyed on relative pronoun "which")
/this purpose/ which is to order the world as an aid to decision and action . . .

IV Discourse Unit 4 (subordinate modifier of Unit 3)

7) Head : The order is what we find to work, conveniently and instructively.
(subordination keyed on semantic repetition of "order" - also operates as recall signal for semantic reiteration of "orderly" from Unit 1)

8a) A. Subordinate modifier of head (keyed on pronoun reference "it"
It is not something we stipulate

8b) B. Coordinate modifier with A (Keyed on repetition of same pronoun reference "it" and repetition of structure "it is not something")
it is not something we can dogmatise about

9a) C. Coordinate modifier with A (keyed on repetition of same pronoun reference "it")
It is what we find ;

9b) 1. Subordinate modifier of C (keyed on repetition of same pronoun reference "it" and repetition of structure "it is what we find") :
it is what we find useful

FIGURE 11

KAPLAN'S ANALYSIS : THE SOIL EROSION PASSAGE

DISCOURSE BLOC

I Discourse Unit 1 (prefatory statement / topic sentence)

1) Every geologist is familiar with the erosion cycle.

II Discourse Unit 2 (subordinate modifier of Unit 1 keyed on "erosion cycle/erosive forces")

2) Head No sooner has an area been raised above sea level than it becomes subject to the erosive forces of nature.

3) A. Subordinate modifier of head (keyed on "erosive forces - rain . . frost" etc as specifications of general)
The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets and then into rivers and out to sea.

4) 1. Coordinate modifier with A (keyed on equivalence of "rain", "frost", etc)
The frost freezes the rain water in cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust.

5) 2. Coordinate modifier with A
Blocks of rock dislodged at high levels are brought down by the force of gravity.

6) 3. Coordinate modifier with A
Alternate heating and cooling of bare rock surfaces causes their disintegration

7) 4. Coordinate modifier with A
In the arid regions of the world the wind is a powerful force in removing material from one area to another.

8) B. Subordinate modifier of head (keyed on "erosive forces" ("rain ... frost ... force of gravity ... alternate heating and cooling ... the wind") /"all this"
All this is natural

III Discourse Unit 3 (coordinate with Unit 2, keyed on contrast ("but") and "erosive forces . . . defensive forces")

9) Head But Nature has also provided certain defensive forces

10) A. Subordinate modifier of head (keyed on "defensive forces / protected")
Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks.

11) B. Subordinate modifier of head (keyed on "soil/different types of soil)
Slowly but surely, different types of soil,with differing profiles, evolve, the main types depending primarily on the climate.

IV Discourse Unit 4 (subordinate modifyer of Unit 3, keyed on "protected by soil/protective soil covering")

12) Head The protective soil covering, once it is formed, is held together by the vegetation.

13) A. Subordinate modifier of head (keyed on "vegetation/grass and herbaceous plants")
Grass and herbaceous plants, with long, branching tenuous roots, hold firmly together the surface particles.

14) B. Coordinate with head (keyed on "the same")
The same is true with the forest cover.

15) 1. Subordinate modifier of B (keyed on "forest cover/the giant trees")
The heaviest tropical downpours, beating on the leaves of the giant trees, reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil.

V Discourse Unit 5 (subordinate modifier of Unit 4 keyed on "held together by the vegetation/thus protected by grass, herbs and trees")

16a) Head The soil, thus protected by grass, herbs and trees, furnishes a quiet habitat for a myriad varied organisms.

16b) A. Subordinate modifier of head (keyed on "a myriad varied organisms/earthworms. . bacteria") as specifications of the general)
earthworms that importantly modify the soil

16c) 1. Coordinate modifier with A (keyed on equivalence of "earthworms/bacteria")
bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants

VI Discourse Unit 6 (coordinate to Unit 5 (keyed on assumption that chemical action, like bacterial action, takes place after the soil has become stabilised and protected)

17a) Head Chemical action is constantly taking place

17b) A. Subordinate modifier of head (keyed on "chemical action/soil acids attack . . . salts in solution move" as specifications of general)
soil acids attack mineral particles

17c) 1. Coordinate modifier with A (keyed on equivalence of "soil acids/salts in solution")
and salts in solution move from one layer in the soil to another

VII Discourse Unit 7 (Subordinate to ?, keyed on "we may sum up" : a generalisation)

18) Head We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions and the parent rock substances develops

19a) A. Subordinate modifier of head (keyed on "it")
It is a long process

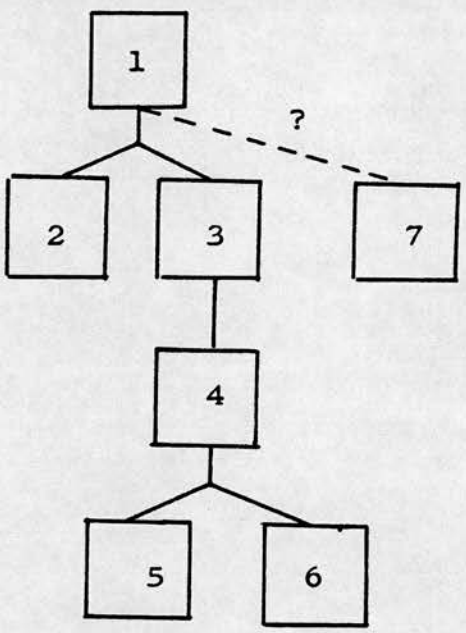
19b) 1. Coordinate modifier with A (keyed on "process/completion of the process")
and in many areas there has not been sufficient time in the geological sense for the completion of the process

19c) i) Subordinate modifier of A and 1 as a coordinated pair, keyed on assumption that i) is result of A and 1)
the soils are immature.

FIGURE 12

SIMPLIFIED TREE DIAGRAM OF KAPLAN'S ANALYSIS

(THE SOIL EROSION PASSAGE)



(Each box represents a discourse unit)

be sub-divided into smaller units dependent upon it, while not necessarily being in its turn dependent on some higher unit.

Kaplan illustrates his ideas with an analysis of the same passage already used by Christensen and Pitkins. (Figure 10). His analysis is visually similar to Christensen's, but he treats the whole passage (or paragraph) as a discourse bloc, and sub-divides it into lower level blocs, or discourse units. This allows the hierarchical structure of the passage to be better represented, as instead of each individual sentence being related to that preceding it, units are related to units. His labelling is somewhat different from Pitkins', being what can only be called a type of structural-semantic parsing. Although he emphasises the importance of performing a functional analysis, he concentrates on cohesive signals rather than actual functions as justification for his divisions.

5.2 Evaluation The soil erosion passage has been analysed for a third time to illustrate Kaplan's technique. (Figure 11) Due to its length analysis stops at sentence level, although Kaplan in his example goes down to clause level (eg Figure 10).

The choice of boundaries for discourse units is partly arbitrary. Sentence 1 has been treated as a separate unit (as in Figure 10), but sentence 9 could also have been so isolated on the grounds that it too acts as an important "head" (Kaplan's term), dominating several recognisable units or blocks. Presumably, the higher up in the hierarchy the sentence appears the more likely it is to be designated as a separate discourse unit.

The present analyst chose to isolate seven separate units, but this was an arbitrary and subjective choice, as more, or fewer, could have been selected. For example, a case could be made for sentences 14 and 15 as a separate unit, and the amalgamation of sentences 5 and 6 is also possible. The summary section is one more difficult to deal with, as we have to decide to which unit it is subordinate. It is clearly at a higher level of generality than what immediately precedes it, but if it is labelled as superordinate to Unit 4 it is not clear where it should be placed in relation to Units 1, 2 and 3. This difficulty comes clearer if we draw a simple tree diagram of the relationship between the seven units (Figure 12).

Pitkins' model is more streamlined than Kaplan's in that he makes explicit the inter-relationships present, leaving the interpretation open. From his diagram it is possible to read off the boundaries of any number of lower level discourse blocs, depending on the level of the text under examination, but no particular choice need be made by the analyses.

Although Kaplan aims to improve on Pitkins' model, he mainly succeeds in confusing what was basically simple. It is doubtful, also, whether the students for whom he intends this method of analysis would find it any simpler to handle than Pitkins' or Christensen's.

6. A.E.Jones and C.W.Faulkner

It is surprising how often work in the same or a closely related field can go on in isolated pockets, simultaneously or in close time proximity, but with no obvious cross-communication. This can happen even within the same

country, and is often the result of a research worker failing to publish in a widely-read journal. Even so, it is a matter for comment that there is no mention in any of the publications so far cited of Jones and Faulkner (1961). It is true that their work appeared as a college composition text book rather than a paper, but as Christensen, Karrfalt, Pitkins, etc are or were all teachers of composition, one would have expected them to have been aware of what other teachers had been doing in the same field. It was left to Vivian Horn (1968) to attempt to marry Christensen's analysis, based purely on relationships of coordination and subordination, with a modified version of Jones and Faulkner's ideas.

6.1 Jones and Faulkner's model Although Jones and Faulkner preceded Christensen and his followers in time, they are being considered after them on account of the greater comprehensiveness their system displays. Their textbook, nevertheless, is an odd combination of traditional rhetoric and a more modern, linguistically-based approach. They themselves describe their model as "structural", and although they never use the word "hierarchy" their system is well described by the quotation from Pitkins referred to earlier (Section 4).

As we have seen, the analysts previously considered had been feeling their way towards a system of labelling to describe the transitional relationships between units of discourse. In 1967 Larson was advocating the development of a functional analysis in order better to distinguish coordinate from subordinate sentences, and Pitkins a year later was also suggesting a finite catalogue of functions. Jones and Faulkner,

however, had already taken such a step. Six years earlier they had suggested nineteen different meaning relationships in four major classes (Table 4), which could be used to link all levels of discourse - clauses, clusters, paragraphs and themes (a "cluster" being a group of sentences forming a unit at below paragraph level). A similarity is evident between Jones and Faulkner's four classes and Pitkins' relationships of coordination, subordination, complementation and superordination. These meaning relationships can be used to produce higher level units which fall into five classes, shown in Table 5.

Jones and Faulkner are primarily concerned with synthesis, but their system can equally well be used for analysis. Five short and simple "paragraphs" corresponding to the five classes in Table 5 are given, to illustrate how the system works (Figure 13). It should be remembered that this system is supposed to be equally applicable to levels both above and below the "paragraph".

6.2 Evaluation Jones and Faulkner differ from all the other analysts considered in using non-authentic texts of a highly artificial type. Their examples (eg in Figure 13) are often suggestive of the stilted thin-textured writing, with an over-use of explicit signalling, common among second language learners. This is unfortunate if their intention is to improve the writing skills of their readers, and lays it open to doubt whether such a tightly structured method of analysis could in fact be used on more complex authentic texts.

TABLE 4

JONES AND FAULKNER'S MEANING RELATIONSHIPS

ENUMERATIVE	EQUAL MEANING	SUBSIDIARY MEANING	DOMINANT MEANING
1. Related action	1. Contrast	1. Definition	1. Generalisation
2. Parallel idea	2. Alternative	2. Amplification	2. Inference
	3. Balanced comparison	3. Sample item	
	4. Result	4. Listing of all items	
	5. Question	5. Sample fact	
	6. Answer	6. Listing of all facts	
		7. Supporting data	
		8. Cause	
		9. Analogy	

TABLE 5

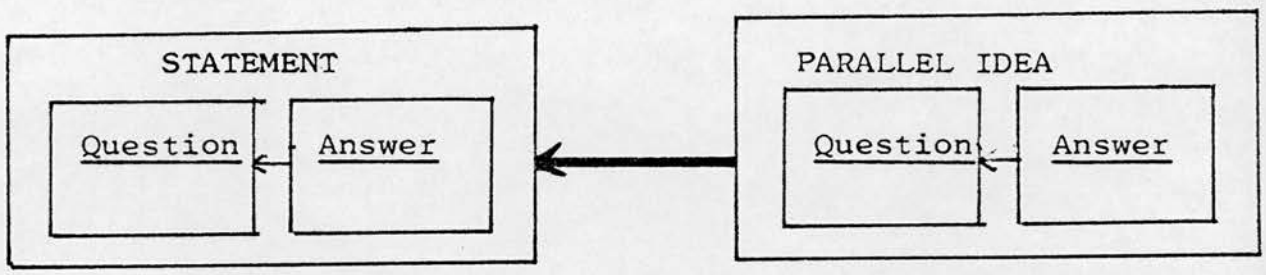
JONES AND FAULKNER'S HIGHER LEVEL CLASSES

CLASS	COMPONENTS
1. Enumeration	Topic Statement + any number of related actions or parallel ideas (= COORDINATION)
2. Equal Pair	Topic Statement + one Equal Meaning Relationship (= COMPLEMENTATION)
3. Unequal Pair	Topic Statement + one Subsidiary Meaning Relationship (= SUBORDINATION)
4. Simple Chain	Topic Statement + two or more Equal Meaning Relationships (= COMPLEMENTATION)
5. Dividing Chain	Topic Statement + at least two Subsidiary Meaning Relationships (= SUBORDINATION)

FIGURE 13

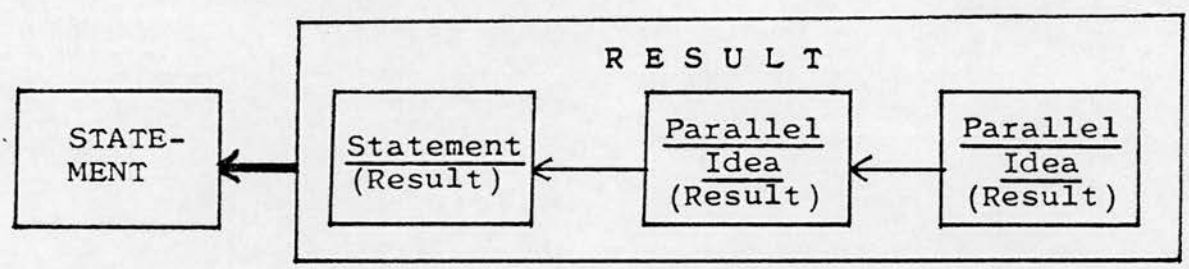
JONES AND FAULKNER'S FIVE TYPES OF PARAGRAPH

1. The Enumerative Paragraph



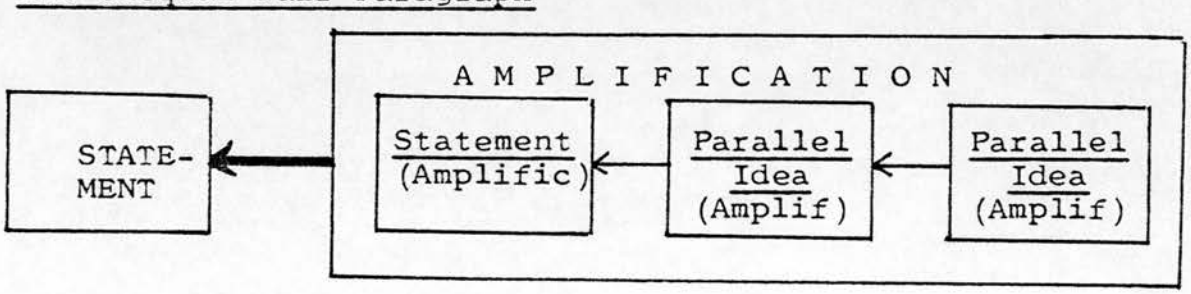
Can the nations of the world - including Russia and China - agree upon a programme of nuclear disarmament? One would hope so, but the indications so far seem unpromising. Can the human race, then, avoid a full-scale, all-out nuclear war if the missile race continues? This seems not only unlikely, but almost impossible.

2. The Equal Pair Paragraph



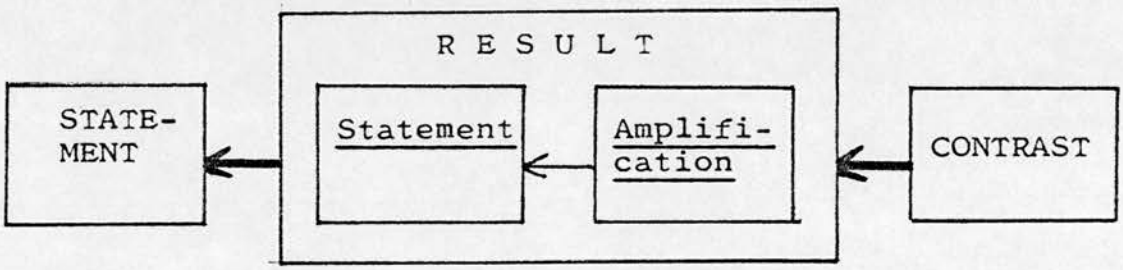
My brother bought a second-hand convertible just before returning to college last September. The first result was that he immediately became popular with a lot of girls, since coeds prefer to date fellows with automobiles. The second result was that he was soon broke, since girls and cars are expensive luxuries. The third result was that he eventually found himself on academic probation, since he had been neglecting his studies during his wild pursuit of happiness.

3. The Unequal Pair Paragraph



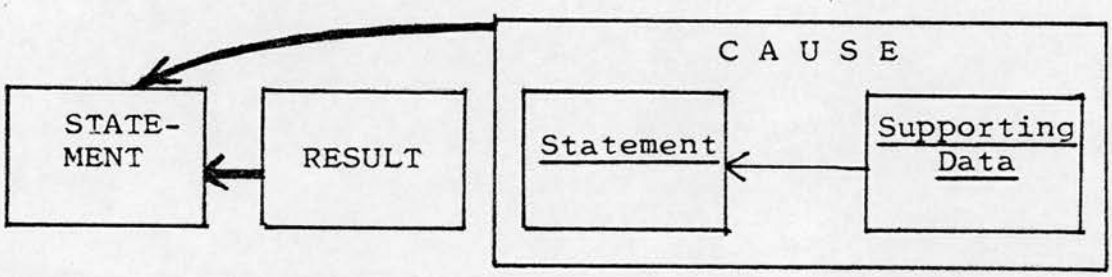
Although it has the same formula as ordinary water (H2O) "polywater" (polymerized water) has some very unusual characteristics. In the first place, it will boil only at a temperature of 1000 F. In the second place, it will not evaporate, even though it is exposed to the air for a long time. In the third place it will not freeze - although at about -40F with little or no expansion it hardens into a glassy substance quite unlike ice.

4. The Simple Chain Paragraph



Smugly confident that he was capable of producing a journalistic masterpiece, the cub reporter became careless and neglected to check his facts. As a result his news-story contained a serious error - he had listed the wrong man as a murderer. Fortunately, however, the city editor discovered the mistake before it got into print.

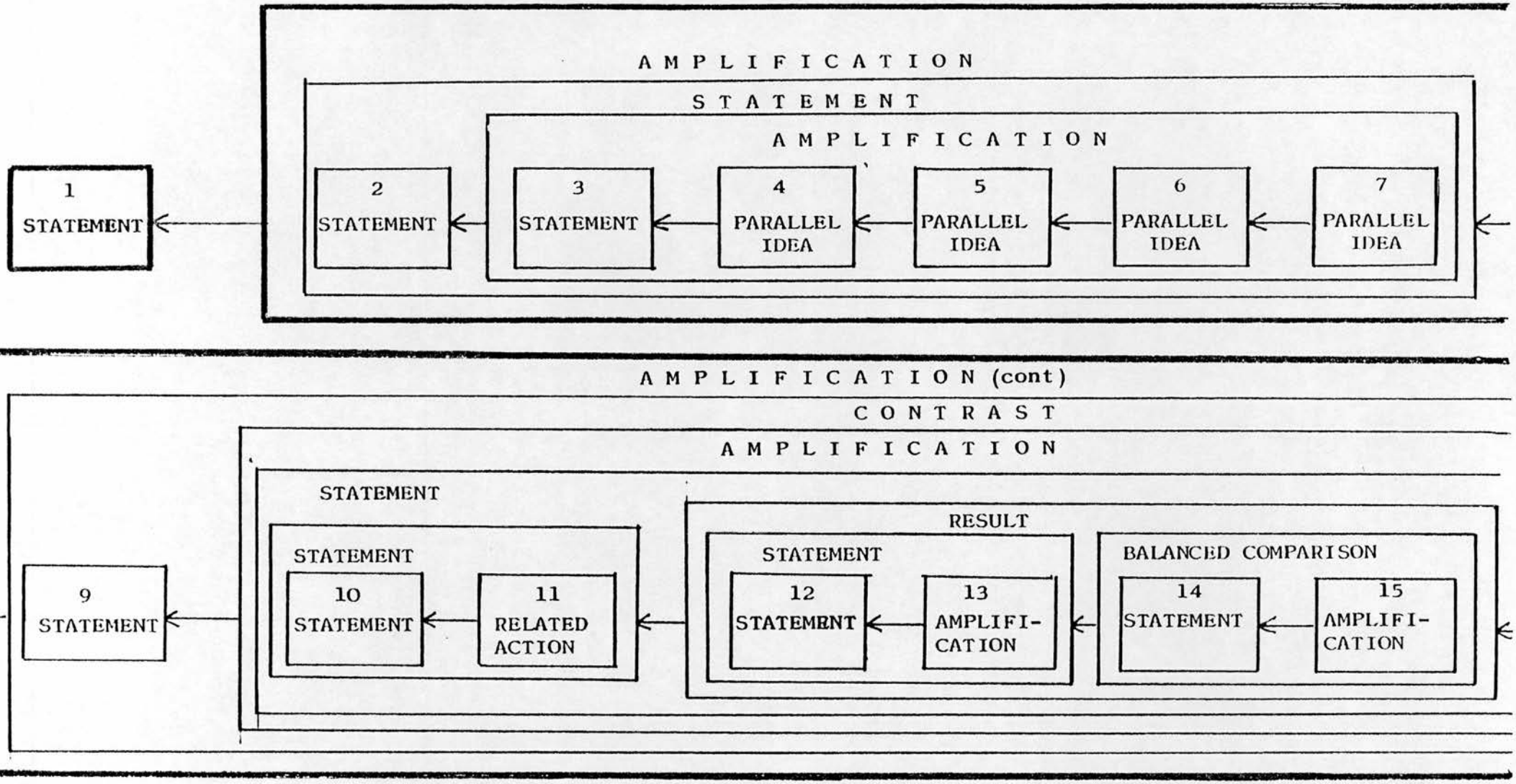
5. The Dividing Chain Paragraph



My cousin has gained nine pounds in four weeks. As you might guess, she now finds her dresses too snug. In my opinion, her increase in weight has been caused by her having ignored her diet and treated herself to candy bars, chocolate malts, and other between-meal snacks. I've seen her munch away through a whole package of cookies while reading a magazine.

An attempt has been made to do just this with the soil erosion passage. (Figure 14). Before going on to discuss specific difficulties, however, it should be noted that this method of diagrammatic representation has some advantages over others so far considered. Pitkins' model, while neat, is essentially static. The text is frozen and analytically dissected, concealing the essential forward movement of discourse. Pitkins is reacting against the traditional approach, which he castigates as ". . . the two-dimensional train, with its topic sentence or introductory paragraph engine pulling a string of freight-car sentences, each with its individual semantic load". (Pitkins 1969 : 140). In his reaction, however, he over-emphasizes hierarchical organisation at the expense of linear organisation. Jones and Faulkner's model, in contrast, represents the constant ebb and flow forward to the next sentence, and then backward to preceding sentence or discourse block, which is the key feature of Christensen's principle of Movement, while at the same time symbolising the hierarchical structure by means of boxes within boxes. Discourse is seen both as a hierarchy and a string.

Despite the very different representations, however, Jones and Faulkner's model is essentially the same as that of Pitkins, and has similar advantages and disadvantages vis a vis the others. Both are based on immediate constituent analysis, and this is made clearer by the conversion of both models into tree diagrams (Figures 15 and 16). Both have the disadvantage of not showing graphically which sentences are subordinate to which, but this problem has been met by Pitkins' use of prosodic feet, and Jones and Faulkner's labelling of meaning relationships.


AMPLIFICATION
STATEMENT
AMPLIFICATION
1 STATEMENT
2 STATEMENT
3 STATEMENT
4 PARALLEL IDEA
5 PARALLEL IDEA
6 PARALLEL IDEA
7 PARALLEL IDEA
AMPLIFICATION (cont)
CONTRAST
AMPLIFICATION
STATEMENT
STATEMENT
10 STATEMENT
11 RELATED ACTION
RESULT
STATEMENT
12 STATEMENT
13 AMPLIFI-CATION
BALANCED COMPARISON
14 STATEMENT
15 AMPLIFI-CATION
9 STATEMENT

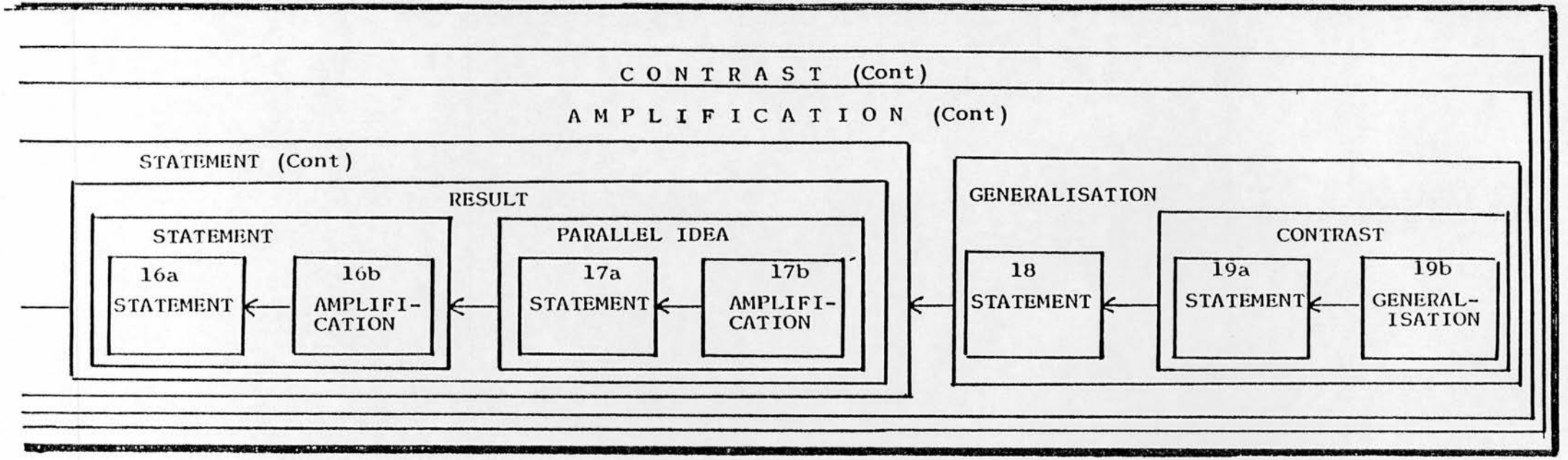


FIGURE 14. JONES AND FAULKNER'S ANALYSIS

THE SOIL EROSION PASSAGE

FIGURE 15

HIERARCHICAL TREE OF PITKINS' MODEL : THE SOIL EROSION PASSAGE

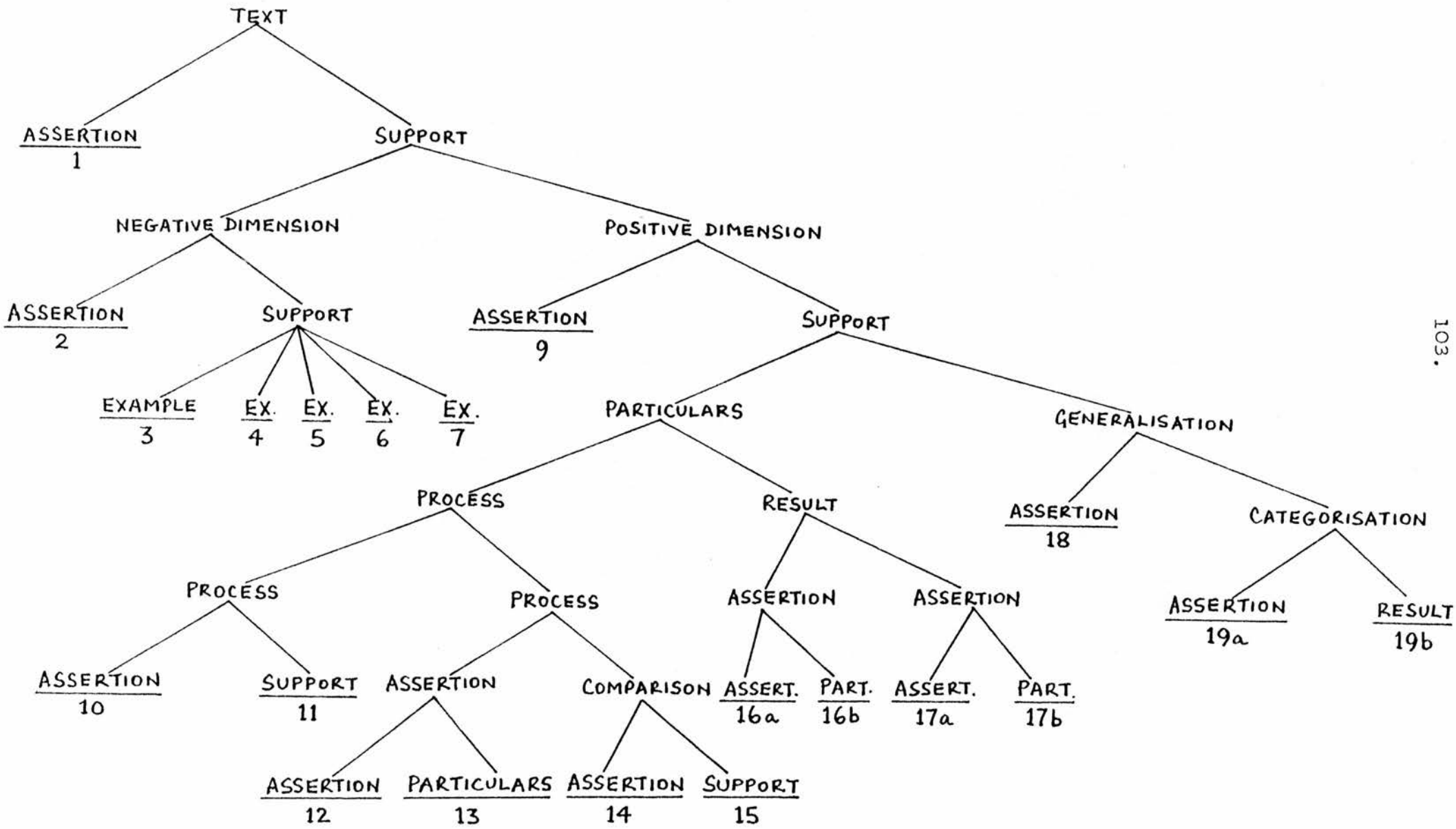
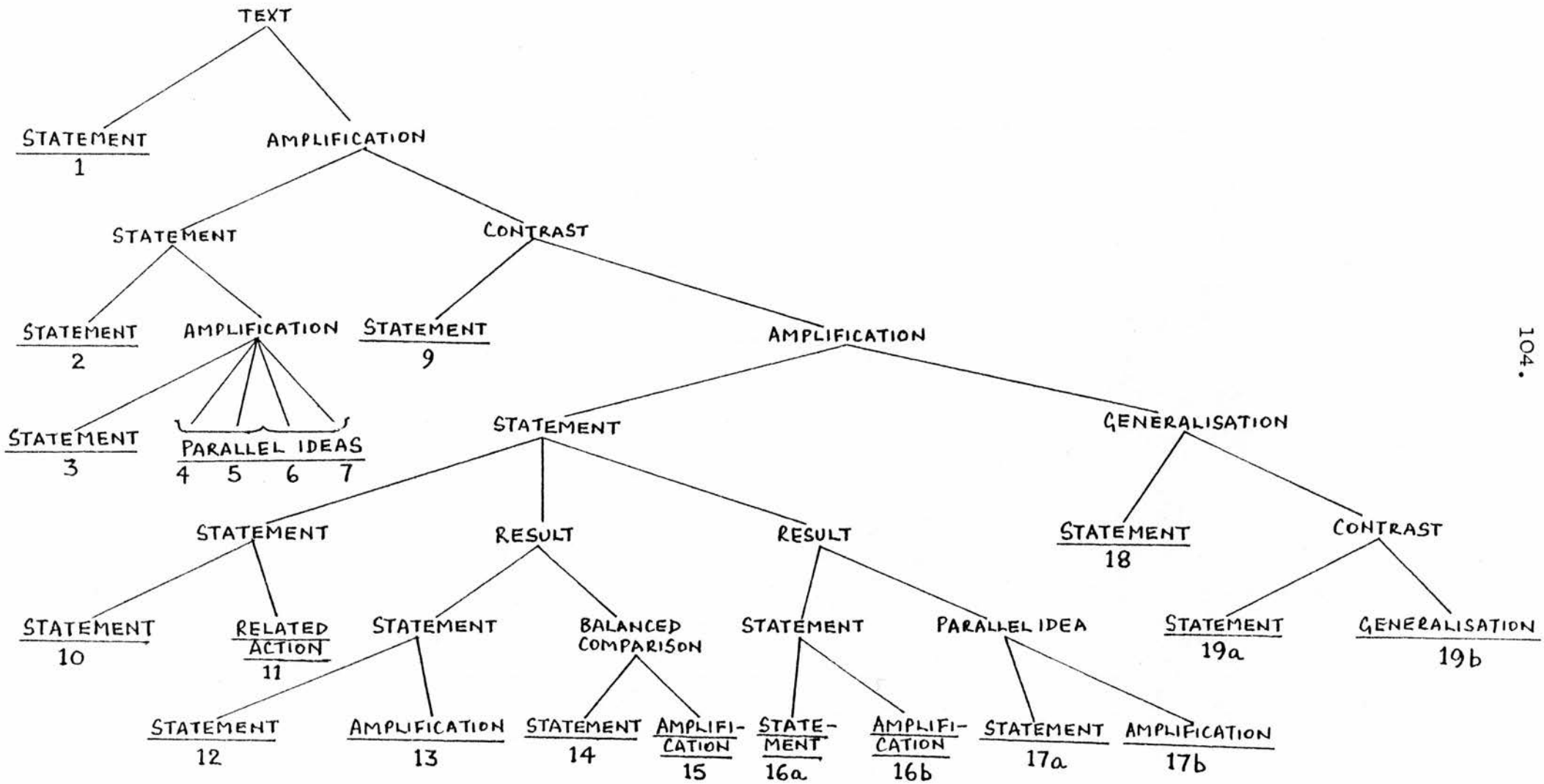


FIGURE 16

HIERARCHICAL TREE OF JONES & FAULKNER'S MODEL : THE SOIL EROSION PASSAGE



Pitkins' model is more rigid here, however, as of his four possible relationships three are binary, only coordination being an exception. Jones and Faulkner's tree is therefore marginally simpler as a result of more flexibility over binary division. Thus, one level fewer is needed by using a simple chain paragraph for sentences 10 to 17.

Most of the problems presented by the Christensen analysis of the soil erosion passage have been eliminated, but not that of sentence 8. It should be remembered that two solutions have so far been put forward. In Christensen's and Kaplan's models "All this is natural" was considered as subordinate to the preceding sentence (Christensen) and to sentence 2 (Kaplan), on the grounds that "all this" refers to the preceding "erosive forces of nature". In Pitkins' model, in contrast, it was presented as a bridging sentence between the negative and positive dimensions, as a marker of complementation at a higher level which might pass unnoticed without a positive signal. In Jones and Faulkners' analysis, however, sentence 8 has been omitted. In a more tightly formulated system no room is left for bridging sentences. This is unfortunate, as the bridging sentence is a fairly common phenomenon of expository discourse (See Chapter 7).

7. Alton Becker

Becker's approach, like Christensen's, was paragraph-based. Instead of looking at the structure of the total text he took individual paragraphs, and attempted to produce a model for paragraph structure which would be generally applicable. Becker and Christensen became engaged in a continuing debate in the pages of College Composition and Communication, and were joined in this by Paul Rodgers, who displayed the cynicism of the literary scholar towards any attempt to structure language. Their debate was reprinted in 1966 as a special edition of College Composition, entitled Symposium on the Paragraph, and again by the National Council of Teachers of English (NCTE) as The Sentence and the Paragraph.

7.1 Becker's theory of paragraph analysis. Becker, influenced by tagmemic theory, saw the paragraph as displaying trimodality. First of all, it consisted of separate functional parts (the feature mode viewed from the particle standpoint). Then, it diaplayed continuity between its parts (the manifestation mode, viewed from the field standpoint). Finally, the whole paragraph was a system of semantic relationships (the distribution mode, also viewed from the field standpoint). Although his analysis neglected the paragraph as a wave, it

is important that he at least recognised it as a dynamic process as well as a static entity.

After examining a large number of expository paragraphs, Becker claimed that two main patterns predominated. These he termed as

1) $+T \pm R + I_n$

2) +Q +A (Originally +P +S)

For each slot there exists a set of fillers. The T (topic) slot is usually filled by a simple proposition. The R (restriction) slot may contain a restatement of T, a definition of T, or an expansion of T. The I (illustration) slot has a variety of manifestations : it can be filled by an illustration, an example, an extended analogy, a series of comparisons, etc. The Q slot contains a question or the statement of a problem, and the A (answer) slot very often has a TRI pattern embedded within it.

Becker gives the paragraph reproduced in Figure 17 as an example of the TRI pattern (Becker, 1965, Young and Becker, 1966), and that in Figure 18 as an example of the QA pattern with embedded TRI (Young and Becker, 1966).

7.2. Evaluation In trying to establish patterns which have universal application for English expository prose, Becker was forced by the variety of actual paragraph structure he encountered to produce such a simplistic model that it can hardly be considered of much value. As a description of

FIGURE 17

BECKER'S ANALYSIS (1) THE TRI PATTERN

(T) [1]The English Constitution - that indescribable entity - is a living thing, growing with the growth of men, and assuming ever-varying forms in accordance with the subtle and complex laws of human character. (R) [2]It is the child of wisdom and chance. (I) [3]The wise men of 1688 moulded it into the shape we know, but the chance that George I could not speak English gave it one of its essential peculiarities - the system of a cabinet independent of the Crown and subordinate to the Prime Minister. [4]The wisdom of Lord Grey saved it from petrifaction, and set it upon the path of democracy. [5]Then chance intervened once more. [6]A female sovereign happened to marry an able and pertinaceous man, and it seemed likely that an element that had been quiescent within it for years - the element of irresponsible administrative power - was about to become its predominant characteristic and change completely the direction of its growth. [7]But what chance gave, chance took away. [8]The Consort perished in his prime, and the English Constitution, dropping the dead limb with hardly a tremor, continued its mysterious life as if he had never been.

(From Lytton Strachey Queen Victoria)

FIGURE 18

BECKER'S ANALYSIS (2) THE QA PATTERN

(Q) [1]Is the United States a nation composed chiefly of people who have not grown up, who think and act with the impulsiveness of adolescents? (A-T) [2]Many shrewd observers of the American scene, both abroad and here at home, are saying that this is indeed the case. (R) [3]They intentionally disturb our patriotic complacency. (I) [4]They bid us view with alarm cultural immaturity revealed by current trends in journalism, by the radio, by the motion picture, by magazines and best-selling books, by mass response to emotionalised propaganda - political and otherwise; by a patent decay of good manners, by the spread of divorce and by other manifestations of parental irresponsibility ; by all the various aspects of behaviour which indicate to a student of human affairs the health or sickness of a civilization.

(From an article by Bernard Iddings Bell in New York Times Magazine 1947)

paragraph development it tells us no more than Genung's 1893 formula (See Chapter 2, 3.5). Becker's Topic (T) is surely the same entity as Genung's 'subject proposed', Becker's Restriction (R) equivalent to Genung's 'whatever is needed to explain the subject', and his Illustration (I) equivalent to 'whatever is needed to establish the subject'. It is true that Genung lacks a theory to support his model, but Becker's tagmemic theory, while valididating his patterns, does not make them any more revealing, or useful for the teacher of composition.

A serious problem is the breadth of the terms used, a point taken up in detail by Rodgers (1966). Rodgers points out that there is very little difference made between the T and Q slots, and makes the following telling observation, which is a criticism as much of Christensen and all the traditional rhetoricians as of Becker :

> The statement of topic has never, to my knowledge, been defined thoroughly. Until we define it, we shall not be able to describe it to everyone's satisfaction, or even to agree about its location.
>
> (Rodgers 1966 : 66)

Rodgers particularly criticises the vagueness of definition of the R slot. He points out that important differences result, depending on the kind of relationship the filler for R has to the filler for T. If the topic is narrowed, as Becker suggests, then the scope of the T assertion is restricted. If it is defined, then the scope remains the same. If, however, it is expanded, the scope of the assertion is widened. Becker also suggests that R

could be a restatement : in this case the topic has to be rewritten, so will be inevitably somewhat changed. Rodgers asks whether it is reasonable to put all these possibilities into the same slot and give them the same label.

The unworkability of the model becomes obvious when we attempt to fit it to actual texts. The two patterns obviously do occur, but not by any means as regularly as Becker would have us believe. In fairness, however, he does point out that these patterns only apply to expository paragraphs - other patterns, on which he does not elaborate, are found in narrative, descriptive and argumentative paragraphs. Expository prose does not only consist of expository paragraphs, and those paragraphs which do not fit the patterns may not in fact be expository at all. Becker also recognises transitional paragraphs (which would include introductions and conclusions) and badly constructed paragraphs as among those to which his patterns would not apply.

An analysis of the soil erosion passage according to Becker's model will clarify some of these difficulties (See Figure 19). It seems impossible to deal with this passage without modifying Becker's formulae. The main problem is the paragraph's hierarchical structure. Sentence 1 is an overall topic statement, but sentences 2 and 9 are also topic statements, albeit at a lower level. One could interpret these lower level topics as restrictions on sentence 1, as has in fact been done in Figure 19, in that they are expansions of it, but according to Becker's theory there can only be one R slot following a T slot.

FIGURE 19

BECKER'S ANALYSIS : THE SOIL EROSION PASSAGE

(T) [1]Every geologist is familiar with the erosion cycle. (R_1) [2]No sooner has an area been raised above sea-level than it becomes subject to the erosive forces of nature. (I) [3]The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets, and then into rivers and out to sea. [4]The frost freezes the rain water in cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust. [5]Blocks of rock dislodged at high levels are brought down by the force of gravity. [6]Alternate heating and cooling of bare rock surfaces causes their disintegration. [7]In the arid regions of the world the wind is a powerful force in removing material from one area to another. [8]All this is natural. (R_2) [9]But Nature has also provided certain defensive forces. ($I\text{-}T_1$) [10]Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks. [11]Slowly but surely different types of soil with differing profiles evolve, the main types depending primarily on the climate. (T_2) [12]The protective soil covering, once it is formed, is held together by the vegetation. (I) [13]Grass and herbaceous plants, with long, branching, tenuous roots, hold firmly together the surface particles. (T_3) [14]The same is true with the forest cover. (I) [15]The heaviest tropical downpours, beating on the leaves of the giant trees,

reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil. (T_4) [16a]The soil, thus protected by grass, herbs and trees, furnishes a quiet habitat for a myriad varied organisms : (I) [16b]earthworms, that importantly modify the soil, bacteria,active in their work of converting falling leaves and decaying vegetation into humus and food for the growing plants. (T_5) [17a]Chemical action is constantly taking place : (I) [17b]Soil acids attack mineral particles and salts in solution move from one layer in the soil to another. Unanalysed [18]We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions and parent rock substances develops. [19a]It is a long process, and in many areas there has not been sufficient time, in the geological sense, for the completion of the process : [19b] the soils are immature.

An alternative interpretation would be to label sentence 1 as T and the rest of the text as I with embedded TRI patterns. The problem here is that Becker does not allow for embedding within the I slot, only within the A slot.

A third alternative would be to label sentence 1 as Q. On the surface sentence 1 is not a question, but it can be argued that a reader comes to expository prose expecting answers. When he reads the introductory statement, 'Every geologist is familiar with the erosion cycle', his expectation is that his own internal question stimulated by this statement (ie 'What is the erosion cycle?') will be clarified by what follows. (See Gray 1977 for an unusual treatment of this point). Thus, the statement itself could be considered as a question in the rhetorical deep structure. Here Rodgers' criticism, that Q is only a special form of T, and therefore need not be differentiated, is particularly relevant. One could extend this to saying that in that case A is only a special form of I, and there therefore seems no logical reason why I should not contain embedded TRI patterns.

Most of the passage divides easily into a succession of TI patterns, but further problems arise from sentence 9 onwards. For example, it is difficult to identify sentences 10 and 11. They have been analysed in Figure 19 as together forming T_1, but it could just as easily be claimed that sentence 11 is an R, with the preceding sentence as T. This again highlights the problem of R - even Becker is not clear

as to its exact status. It was decided not to identify an R slot here because Becker does not allow for TR without I, but this still leaves the problem of an isolated T.

From sentence 9, also, we have a fairly complicated series of embeddings which do not appear as such in Becker's analysis. He does, somewhat grudgingly, recognise occasional embedding (TRI in an A slot - see Figure 18), but does not allow for the multiple embeddings found in most normal expository prose. Thus the hierarchical structure of the passage is obscured.

The last part of the text, sentences 18 and 19, are left unanalysed, as they fall within Becker's category of transitional paragraph, to which the TRI pattern is inapplicable.

7.3 Field analysis of the paragraph. Becker only briefly touches upon the rhetorical field structure of the paragraph, as, in spite of the tagmemic principle of trimodality, only the feature mode has been fully developed. He claimed, however, that 'it is through field analysis that we begin to understand the organic nature of the paragraph, its ability, like a poem's, to shape itself once its dimensions have been specified'. (Becker 1965 : 242).

A semantic field analysis of a discourse is undertaken in the same way as that of any other unit of language - by means of a matrix. Table 6 illustrates this with an analysis of the I slot of the English Constitution paragraph.

Here the 'dimensions' of the paragraph have been specified by the T and R slots. In these slots we were told

TABLE 6

BECKER'S MODEL : FIELD ANALYSIS OF THE PARAGRAPH

THE I SLOT OF THE ENGLISH CONSTITUTION PARAGRAPH (SEE FIGURE 17)

FORCES SHAPING THE ENGLISH CONSTITUTION

HISTORICAL MANIFESTATIONS	WISDOM	CHANCE
1688	The wise men . . . moulded it into the shape we know.	
(1714)		George I ... gave it ... the system of a Cabinet independent of the Crown and subordinate to the Prime Minister.
(1832)	Lord Grey saved it from petrifaction and set it upon the path of democracy	
(1840)		(Victoria's marriage made it seem likely that a quiescent element) was about to become its predominant characteristic and change... the direction of its growth.
(1861)		(With the death of the Consort) the English constitution . . . continued its mysterious life as if he had never been.
(?)	?	?

that the English Constitution is continually growing according to the wisdom of the men who have influence over it, and the chance events of history. The expectation is raised that the paragraph will continue by stating instances of this 'wisdom' and 'chance', and explaining their effects. The matrix shows us how this development takes place, and also indicates how the paragraph could be extended into present time by adding further, more recent effects.

7.4 Becker's main contributions to the theory of the paragraph. Although Becker's system was little more revealing of paragraph structure, or useful as a teaching tool, than any of the other systems of analysis being advocated in the 1960s, he made a significant contribution in two related areas, the identification of formal markers of text chunking, and the use of experimental procedure for the first time in paragraph analysis.

7.4.1 Formal markers of chunking Formal markers of paragraph tagmemes could, he thought, be classified according to the three tagmemic hierarchies of lexicon, grammar and phonology. These markers occur independently, in their separate hierarchies, but are inter-related, so that where several occur in one place a paragraph division becomes more likely. The three systems will be described with reference to the English constitution passage, as this is a relatively simple passage, which conveniently fits the TRI pattern.

A. Lexical

(i) Lexical equivalence chains (called by Halliday and Hasan (1976) 'cohesive chains') These Becker

considered to be the most important formal paragraph markers. Table 7 shows the two chains evident in the English Constitution passage. The 'English Constitution' chain extends throughout the entire paragraph, while a subordinate chain, that representing the human forces that shaped the Constitution, extends only from sentences 3 onwards. This represents a sub-theme, and identifies the I slot of the paragraph.

(ii) Lexical transitions, placed by Halliday and Hasan (1976) under the general heading of 'conjunction'. In the passage, the only examples are 'then' (sentence 5) and 'but' (sentence 7), in each case introducing a minor chunk within the I slot. Particular transitions may be linked to particular slots. The R slot may be introduced by 'in other words', 'besides', or 'that is', and the I slot by 'for example', or 'for instance'. The continuation of a slot may also be marked by such transitions as 'in addition', 'likewise', or 'furthermore'.

B. Grammatical

(i) Grammatical parallelism This may take the form of parallel sentence structure, as in sentences 1 and 2, and sentences 3 and 4 of the English Constitution passage :

1. The English Constitution is a living thing . . .
2. It is the child of wisdom and chance.

3. The wise men of 1688 moulded it . . .
4. The wisdom of Lord Grey saved it . . .

TABLE 7

LEXICAL EQUIVALENCE CHAINS : THE ENGLISH CONSTITUTION PASSAGE

SENTENCE	MAIN CHAIN	SUBSIDIARY CHAIN
1	The English Constitution = a living thing	
2	the child of wisdom and chance	
3	it, it	The wise men of 1688
4	it, it	Lord Grey
5	↓	↓
6	it	an able and pertinaceous man
7	↓	↓
8	The English Constitution	the Consort

It can also be seen when the grammatical roles of lexical items are parallel. In sentences 1 and 2 'the English Constitution' fills the subject slot, but in sentences 3 and 4 it fills the object slot, thus suggesting a break of some kind between sentences 2 and 3.

(ii) Verb tense A shift in the verb tense may indicate a break. For example, in this passage the movement from the R to the I slot is reflected in a shift from present to past.

C. Phonological

Becker only mentions this in passing, but he suggests that more research is necessary on the shifts in pitch, register, tempo and volume discernable when a text is read aloud. Work done since by Rees and Urquhart (1976) indicates that shift of pitch setting of the first salient syllable of a tone group can be used to indicate discourse chunking.

Becker adds a fourth factor, the graphical, to the above three hierarchies, pointing out that the paragraph is also marked conventionally by indentation.

Although Becker sees the above as signals of paragraph structure, they can be viewed more realistically as markers of text chunking in general. The unity of the English Constitution paragraph as one chunk is only signalled graphically, as without the complete text it is impossible to know where the main lexical equivalence chain begins. Sentences 1 and 2 are marked as a chunk by grammatical

parallelism and change of verb tense, and sentences 3 to 8 by the beginning of the subsidiary lexical equivalence chain and verb tense. We can sub-divide further, however, as sentences 3 and 4 are differentiated by grammatical parallelism, and sentences 5 and 6 and 7 and 8 by lexical transitions. Becker might argue that the TRI pattern itself indicates that this is a paragraph unit, but a TRI pattern may on occasion be embedded in an A slot.

Probably even Becker would not claim that the above is a complete catalogue of chunking signals. More will be written on this point in Chapter 6.

7.4.2 Experiments in paragraph recognition Becker and his associates were convinced of the reality of the paragraph as a patterned unit of language marked by formal cues, and between 1964 and 1968 a series of experiments was conducted on the veracity of this model. (Young & Becker, 1964, Koen, Becker & Young, 1968).

Four expository passages of medium length were chosen, and a second version of each was made, replacing all content words by nonsense words. The authors' paragraph indentations were removed, and extra lines added at the beginning and end of each passage. Then 48 subjects were given one English and one nonsense passage each, and asked to insert the missing paragraph boundaries. On completion of this task, each subject was given a list of seven cues and one open category, and asked to indicate for each boundary one or more of the cues which he had used. The cues were as follows : change of subject, change of time, change of location, change of verb

tense, the beginning or ending of a QA pattern, the beginning or ending of a TRI pattern, and a transition word. The subjects were all students of the experimenters, who were familiar with the rhetorical patterns of TRI and QA.

Before the experiment it was hypothesized that the subjects would divide the passages in approximately the same way, and that there would be a positive correlation between three system breaks (ie. lexical, grammatical and rhetorical) and the subjects' placing of paragraph boundaries.

These hypotheses were upheld by the results. The subjects were generally in agreement, and there was a high correlation between paragraph divisions chosen for the English and nonsense passages. Of twenty junctures at three system breaks, 13 were indicated by more than 20% of the subjects. (If 20% or more chose a particular point, then this was considered significant).

More detailed comments on these experiments will be reserved until the end of Chapter 6, when the results are discussed together with those of a similar experiment carried out by this investigator. Two comments only will be made here. Firstly, Koen, Becker and Young nowhere define exactly what they mean by the term 'paragraph', and it can only be assumed that, as they are concerned with the practical teaching of composition, they are referring to the orthographic paragraph. Secondly,

they consider that a probability of paragraphing of 20% is sufficient for significance, which is surprisingly low, bearing in mind that choice of paragraph boundaries is idiosyncratic to a far greater degree than choice of sentence boundaries. It therefore seems questionable whether quite such definite results can be claimed, when only one subject in five needs to have chosen a boundary for it to be included as significant.

8. Conclusion.

All the writers considered in this chapter, except perhaps for Pitkins, make the error of attempting to fit the paragraph into their system of analysis. Unfortunately for them, in many cases authors do not seem to follow a consistent pattern in paragraphing their work.

Christensen begins with the intention of analysing paragraphs as isolated units, stating that one paragraph ends and the next begins when the relationship of a sentence to that preceding it is neither coordinate nor subordinate. Karrfalt finds that this analysis breaks down when a paragraph ends with a more general or summary statement, and tries to adapt Christensen's model to fit these cases. Pitkins introduces the 'discourse bloc', which Kaplan sees as corresponding to the paragraph, and thus introduces the lower level 'discourse unit'. Jones and Faulkner

see the paragraph as one level in the text hierarchy intermediate between 'cluster' and 'theme', and attempt to classify all levels in the same way. Jones and Faulkners' model seems the tightest and most effective, but they never attempt to apply it to expository text. Had they done so they would certainly have found many awkward exceptions, as did Christensen and Karrfalt.

Becker stands slightly apart from the above group in claiming to base his model on tagmemic theory. The end result, however, is disappointing, and does not take us much further than traditional rhetoric, the sets of slots he suggests only reflecting such formulae as that of Genung (1893). Like Christensen, his model depends to a large extent on the presence of an explicit topic sentence, but unlike Christensen he nowhere attempts a definition of this, so that its identification in actual text is even more problematic than it is in Christensen's model.

The frequent lack of correlation between an author's chosen paragraph pattern and the analyst's expectations is partly the result of these writers never making completely clear what type of prose they are considering. Even within expository prose there are certainly occasional non-expository paragraphs which might well follow a different pattern, and in any case none of these writers actually defines what he means by expository prose.

This problem might be partially resolved if a distinction were recognised between the orthographic and the conceptual paragraphs (an idea put forward by Lackstrom, Selinker and Trimble (1973) and later revised by Urquhart and Widdowson (1976) and Rees and Urquhart (1976)). The orthographic paragraph is a physical unit marked by a new line and an indentation. The conceptual paragraph is a discourse unit, and is a chunk of text which hangs together on account of its internal relationships. The two usually coincide but need not do so; an orthographic paragraph may correspond exactly to a conceptual paragraph or it may contain two, three or even more conceptual paragraphs within it, this depending on such extra-textual factors as audience, subject-matter, format, or even the idiosyncratic preference of the writer. Pitkins' "boxes within boxes" can thus easily be related to possible patterns of paragraphing. These will be dealt with in a more detailed way in the experimental section of this work.

The special value of Becker's work lies not so much in his model as in his suggestion that paragraphs themselves, or their constituent 'chunks', can be identified by formal markers, which he classifies according to the three tagmemic hierarchies of lexicon, grammar and phonology. Together with Koen and Young

he was also the first to carry out experimentation on paragraph recognition (Young and Becker (1964), Koen, Becker and Young (1968)).

CHAPTER 4

TOWARDS A THEORY OF THE PARAGRAPH

1. Later developments and their limitations.

Concern with the phenomenon of paragraphing seems to have lessened in recent years. Of the relevant areas of enquiry prominent in the 1970's, namely Speech Act Theory and Cognitive Science, the former follows on more directly from Traditional Rhetoric than does the latter. Speech Act Theory, however, offers us very little as yet to explain paragraphing, and has thus been omitted from consideration here.

Cognitive Science, on the other hand, does suggest a new approach, an approach which arises directly from the study of perception and human information processing procedures. A few psycho-linguistic experiments have indirectly confirmed some previously unsupported statements made by traditional rhetoricians (eg. Bransford and Johnson (1972), Kieras (1978)). Of more significance, the concept of the frame (Minsky (1975), Winograd (1977)) places the structuring of expository prose within a more global view of perception in general, and justifies the construction of paragraph models despite the lack of correspondence found between such models and many

authentic paragraphs. Moreover, Van Dijk's (1977) theory of the macro-structure throws some light on the topic sentence, and could form the basis of a theory of summary.

A possible future approach might be to consider text structure from a perceptual standpoint. The paragraph, with its more general topic sentence and more specific supporting sentences, could be seen as the most efficient means by which the human mind is able to process and internalise the salient points of an extended written discourse. Psycho-linguistic research, however, is a relatively new field, and for the moment too little is known about the workings of the human mind to build up a viable theory.

2. The paragraph in traditional rhetoric.

To return to the basic concern of this work, the survey in the two preceding chapters should have made it clear that traditional rhetoric can provide us with a rich source of information on the structure and paragraphing of expository prose. In particular, such early writers as Genung had considerable insight into the subject, insight which can still be of major importance both in the classroom and in research. Unfortunately, however, these early writings suffer from the same defects as do those of the traditional grammarians - they rely on

piecemeal and unsystematised material, and present no satisfactory unified theory.

In spite of these drawbacks, the rhetoricians' concept of the paragraph is that which persists in the mind of today's educated but linguistically unsophisticated layman. Whether this is the result of generations of unchanging school instruction or of an inherent intuitive reality is impossible to say. In the words of H.G. Widdowson, however, 'the language user's intuitive sense of the nature of language as expressed through these much maligned definitions has its own legitimacy, and we would do well to respect it.' (Widdowson, 1979 : 237) Widdowson was referring here to the traditional definitions of the parts of speech, but his words could apply equally well to traditional rhetoric as to grammar.

Among the endless traditional prescriptions for paragraph construction with which the student has been assailed over the last century the most important are the principles of Unity and Continuity. 'Unity' necessitates a topic sentence, usually explicit but sometimes implicit, stating the subject of the paragraph and supported by sentences bearing some relation to that subject. 'Continuity' (or 'Coherence') is obtained by making sure that the theme is passed from sentence to sentence by means of transitional links. Apart from these areas, traditional rhetoric

has been a mine of information for the student on such topics as sentence length, paragraph length and the internal organisation and composition of the paragraph.

Little has been added to such basic concepts by the later workers of different traditions, but some suggestions have been made as to how so many loose ends can be tied up to form a unified whole. Most work directly on paragraphing was done in the 1960s by American teachers of expository writing, including some who were trained as tagmemicists, their main contribution being the idea that the organisation of expository prose is hierarchical. Christensen pursued the relationship between sentence and paragraph to suggest a theory of paragraph unity, ie that the topic sentence is the sentence upon which the remaining sentences of the paragraph depend, in a relationship of coordination and/or subordination. Christensen himself never went beyond the single paragraph, but his pupils and successors moved on to deal with whole texts, Pitkins being the first to make the hierarchical nature of prose structure explicit.

Becker and his colleagues developed the idea of formal signals as markers of movement from one information block to another, and were the first to use experimental procedures to investigate

paragraphing. Like Christensen, however, Becker made the mistake of pursuing an exclusively surface approach, and as a result failed to recognise the difference between the orthographic paragraph and the conceptual paragraph (or information block).

An overall picture begins to emerge from the above survey, but a basic confusion persists throughout, a confusion which needs to be resolved if a theory of expository prose structure in relation specifically to paragraphing is to result. This is the confusion between the paragraph as a discourse unit (the information block) and the paragraph as an orthographic unit. Pike and Pike realise this distinction when they define what they label as 'a paragraph', but what is termed here as an 'information block' as 'a theme stated and developed' (Pike and Pike, 1977 : 25), and separate it off from the orthographic paragraph, which they prefer to call a 'sentence cluster'. Lackstrom, Selinker and Trimble refer to the same features as the 'conceptual paragraph' and the 'orthographic paragraph' respectively. Thus the paragraph is only a surface phenomenon, serving principally to signal the hierarchical organisation of the text, while the information block is the more important discourse unit which, embedded within higher level

information blocks, forms part of the deep structure of the text.

Too often, also, the mistake has been made of considering the paragraph in isolation, when it is in fact only a reality when seen as a chunk of prose within a complete discourse. Its boundaries are chosen by the writer to suit his particular purposes, and it may coincide with one or several information blocks. What happens within the paragraph is not as important as what happens within the information block.

3. A possible model of expository prose structure

From an analysis of the above literature and from some of the personal observations of this writer, the following points are suggested as the basis for a more unified model of expository prose structure and paragraphing.

(i) Expository prose (though not necessarily other kinds of prose, eg. narrative), has a hierarchical structure, consisting of information blocks at different levels from the complete text to below the sentence. Paragraphing is imposed upon this super-structure of blocks within blocks as a surface phenomenon.

(ii) Paragraphs may optionally begin at points where there is a movement upwards to a higher level

in the hierarchy. Whether a new paragraph begins or not depends on a number of external factors, eg the appearance of the paragraph on the page, or the author's wish to isolate a portion of the text for emphasis, but its beginning will always coincide with the beginning of an information block. Paragraphing can be considered as one factor only in indicating the hierarchical organisation of the text and thus facilitating its comprehension.

(iii) The structural organisation of the text is marked by signals, which are sometimes explicit but often only implicit. These mark movement or transition from one information block to another, and the orthographic convention of paragraphing may be considered as one particular type of higher level signal.

(iv) Topic statements, like the information blocks to which they belong, are hierarchically arranged. A topic statement is usually the most general, highest level proposition of the information block, with lower level propositions within the block dependent on it. A topic sentence contains the topic statement of the information block to which it belongs.

(v) Sentence length fulfills an important function in that, if variety of sentence length is followed, the short sentence (ie. short in relation to its neighbours) may serve as a signal for change of topic, ie occur as a topic sentence, though usually at a fairly low level.

(vi) This ideal pattern may occasionally be disturbed by the presence of 'bridge' or 'transitional' sentences, which straddle the boundary between adjacent information blocks. In these cases paragraphing may occur either before or after the bridge.

4. Experimental evidence

The next part of this work will be an account of experiments and text analyses carried out in order to verify some of the above statements. It would, of course, be too ambitious a project to set up and validate a complete model : no more will be done here than to describe four experiments which will hopefully support some of the claims made above, and by means of text analyses illustrate two hypotheses which could form a basis for future research. As the stimulus for this work was a particular pedagogical problem, the orientation of the experiments will be practical as much as theoretical, with the aim not only of

adding to our total understanding of the phenomena of prose structure and paragraphing but also of producing pointers towards some solutions to these problems.

Three authentic texts are used in several places in the investigation (Chapters 6 to 8). These are 700 to 900 word extracts from geography textbooks of the type likely to be read by sixth-formers and first year undergraduates. They were chosen as examples of 'good' expository writing (judged intuitively), which were at the same time structured in such a way as to enable the type of analyses used in the experiments. Any of the social sciences would have been suitable sources, eg. education, sociology, or even linguistics.

The order of treatment will be from the lower level to the higher, ie from the sentence to the complete text, each experiment being to a certain extent the basis for what follows. Although a case can be made for information blocks at below sentence level, the sentence itself will be considered as the lowest level block in this study. The justification for this decision is that the present emphasis is on paragraphing, and paragraphing is of necessity a supra-sentential feature. Introduction to each experiment will be made in the succeeding chapters :

for the moment it is sufficient simply to state the main hypotheses to be tested.

Hypothesis I Experienced readers will tend to reject the extremes of sentence length (ie short and long) in favour of varied and medium-length sentences. There will be a preference for sentences of varied over those of medium length.

Hypothesis II When considering passages consisting of sentences of varied length, experienced readers will tend to prefer those with short sentences at change of topic to those with short sentences randomly placed.

Hypothesis III A paragraph or information block boundary is most likely to occur under either or both of the following conditions:

(a) When there is a movement upwards to a higher node in the hierarchical structure of the text. The greater the vertical distance of the movement, the more likely is it that a new information block boundary or paragraph will begin.

(b) When there is a movement downwards or across the tree to a node which dominates a number of propositions through at least two branches. The more propositions the node dominates, the more likely is it that a new information block boundary or paragraph will begin.

Hypothesis IV Extrinsic signals are more effective markers for the reader of paragraph or information block boundaries than are intrinsic signals.

Hypothesis V Experienced readers are influenced by format when making judgments over paragraphing. They will prefer longer paragraphs when the lines are long and/or single-spaced, and shorter paragraphs when the lines are short and/or double-spaced.

No actual experiments will be carried out to test the reality of the topic sentence or of the bridge sentence, but tentative theories based on evidence presented here will be suggested.

CHAPTER 5

THE SIGNIFICANCE OF SENTENCE LENGTH

1. Introduction

Grimes (1978) makes the following comment about the sentence :

> It is much like a packaging unit in which a certain portion of the content can be arranged for export via sound to the listener. Or, to use another metaphor, we do not expect the hearer to drink our whole kettle of semantic soup at one gigantic gulp ; we give it to him a spoonful at a time.
>
> (Grimes 1978 : 129)

The fact that we are considering discourse does not mean that we can ignore the sentence. The expository writer is inevitably obliged to express his ideas in sentences, and the way in which he arranges his material within sentence boundaries can help or hinder comprehension. Although sentence length is a feature of writing virtually ignored by the modern teacher, even a cursory examination of a number of texts will reveal this as an important stylistic feature of the individual writer, and a factor contributing towards the 'rhythm' and 'balance' recommended by the prescriptions of traditional rhetoric.

The sentence will therefore be the lowest level information block considered in this work, and an investigation will be made into optimum sentence length and the function of the short sentence.

2. Sentence length and the second-language learner

Sentence length is one significant difference between the writing of the educated mother-tongue speaker and that of the intermediate level learner. The learner does not, as one might expect, necessarily write the shorter sentences. This claim is supported by an informal experiment. A representative paragraph was chosen from a geography essay written by a Malawian fourth year secondary school pupil, and seven graduate mother-tongue speakers were presented with the same propositional content and asked to reconstruct the information as a coherent paragraph. In each case the educated mother-tongue speaker wrote fewer words than the school pupil. Five out of the seven wrote as many, or more, sentences, and all wrote sentences which were, on average, shorter than those of the pupil. (See Table 8).

TABLE 8

COMPARISON BETWEEN SENTENCES WRITTEN BY ONE SECOND LANGUAGE AND SEVEN MOTHER TONGUE SPEAKERS

	LEARNER	MOTHER-TONGUE SPEAKERS						
		1	2	3	4	5	6	7
NO. OF WORDS	154	106	139	132	146	118	108	136
NO. OF SENTENCES	5	5	6	7	7	4	4	6
AV. NO. OF WORDS PER SENTENCE	30.8	21.2	23.2	18.8	20.8	29.5	27.0	22.7

This correlates well with Kellog Hunt's research into the development of syntactic maturity in the writing of American high-school pupils. (Hunt 1965, 1970). Among other things, he showed that sentences do not become progressively longer with maturity : in fact, they become longer only up to a certain level, the stage at which excessive coordination is used, and then grow shorter, so that skilled adults write shorter sentences than elementary school-children. Little research has yet been done in this area on the writing of second language learners, but it is reasonable to assume that a similar pattern would be evident.

2.1 Optimum sentence length : the psycholinguistic approach

When we consider the optimum sentence length for a passage of expository prose, two opposing points of view emerge. Grimes' statement quoted above implies that the shorter the sentence, the easier it is for the recipient to comprehend, and this is a belief strongly held by most producers of readability formulas, for whom sentence length is the only widely used non-lexical variable. (See Klare, 1974, for a general survey of such formulas). It is not suggested that the length itself affects comprehensibility, it being more likely that the longer a sentence is the more syntactically complex it will be. According to Gunning :

> The main reason why sentence length is a good measure of reading difficulty is that it measures relationships. The longer the sentences are, the more words ; the more words, the more relationships between them - and consequently the more effort for the reader.
>
> (Gunning 1968 : 57)

It is widely accepted that the more syntactically simple a sentence is, the easier it is to process, with the result that in simplifying texts for second language learners it is usual to reduce subordination and embedding, and rewrite in shorter sentences. Schlessinger (1968) suggests that, apart from their necessarily simple syntax, short sentences are easier to understand than long because they display more redundancy than would be the case with longer sentences, more cohesive features being needed.

A case can be made, however, for the use of long sentences rather than short. The extra redundancy spoken of by Schlessinger may perhaps facilitate comprehension at a very elementary level, at the stage when the reader can only hold a limited amount in his short term memory store, but the ensuing repetition is likely to irritate the more sophisticated reader. More important, though, is the conclusion come to by Bransford and Franks (1972) after an experimental investigation into the reality of the sentence as a unit of memory. They suggest :

> Simple syntax may actually hinder comprehension by forcing Ss to do too much of the integration. And more complex structures (like some forms of embedding) may actually facilitate comprehension by explicitly expressing easily codable semantic integrations of ideas. Indeed, if embedding transformations did not facilitate communication to some extent, it would be difficult to imagine why languages make use of them at all.
>
> (Bransford and Franks 1972 : 245)

The discussion up to this point has mainly concerned comprehension, and it seems obvious that one feature of good

writing is comprehensibility. Our students should, if nothing else, learn to write intelligibly. A balance clearly has to be struck between writing sentences that are so short as to appear irritatingly disjointed, and so long as to be impossible to hold in the short term memory store.

2.2 Optimum sentence length : the language teacher's approach We have already seen in Chapter 2 that the traditional rhetoricians had something to say on this subject, the commonest advice being to avoid monotony by varying sentence length. In Britain in the 1930s, however, rhetoric in the old sense went out of fashion, and when the thread is taken up again we find more and more writers recommending short sentences. This may reflect the influence of the readability formula. For example, Flesch (1949) and Gunning (1968), the originators of popular and easily applicable formulas, both recommend aspiring writers to use short sentences. Gowers (1973, rev. Bruce Fraser), an important reforming influence on civil service writing, suggests :

> Keep your sentences short. This will help both you to think clearly and your correspondent to take your meaning. If you find you have slipped into long ones, split them up.
>
> (Gowers 1973 : 31)

Connors (1974), writing more recently for Open University students, also claims that short sentences are preferable to long ones.

The tradition of old-fashioned rhetoric continued, as we have seen, in the American college composition class, and here we still find variety of sentence length held up as the ideal. As recently as 1976 Irmscher stated :

> The single most important principle of variety in prose style is writing long and short sentences - not long, short, long, short, but long and short as the purpose dictates and as your ear hears them rhythmically. There is no particular stylistic virtue in a long sentence per se, or a short sentence, or one of moderate length. There is virtue, however, in enough variety to avoid the ponderousness of too many long sentences, the choppiness of too many short ones, and the monotony of too many medium-length ones.
>
> (Irmscher 1976 : 106)

Few recent writers have seen any significance in the positioning of short and long sentences. Kaplan (1970) goes as far as to deny any importance at all to sentence length. Donley (1976) notes in passing that short sentences may be significant in indicating sub-topics within paragraphs. Tufte (1971) goes furthest in illuminating us on this subject. She writes :

> The better the writer of fiction or of non-fiction alike, the more he tends to vary his sentence length. And he does it as dramatically as possible . . . . Time and again the shortest sentence in a professional paragraph is brought up against the longest, or at least lodged among some much longer.
>
> (Tufte 1971 : 29-30)

Unfortunately, Tufte limits herself to the consideration of what she calls 'kernel sentences' rather than short sentences as such. It should be obvious, however, that a short sentence is quite likely to be a

kernel, and vice versa, but the significant point is that the sentence is short, not that it is a kernel. She stresses kernels because she seeks to show that such a sentence is basic both grammatically and semantically, and points out six rhetorical functions of the kernel sentence that she has identified. It seems preferable here to consider these as functions of the short sentence rather than the kernel sentence. They are as follows :

1. To act as the topic sentence
2. To indicate a contrast in the middle of a paragraph
3. To indicate a transition between paragraphs, appearing either at the end or at the beginning (compare 1 above)
4. To indicate a negative/positive pattern (compare 2 above)
5. To relieve monotony in the centre of a paragraph between long sentences.
6. For emphasis

It should be noted that apart from points 5 and 6 all these functions relate in some way to movement from one topic to another.

2.3 Definition of 'long' and 'short' As far as the actual length of such sentences is concerned, most writers do not attempt a definition, taking it for granted that a sentence is only short or long relative to other sentences in the text. Wehrlich (1976), however, in a recent text grammar of English, makes the suggestion that in writing the length of a sentence can be measured by the number of printed lines it covers. His short sentence occupies less than one completed printed

line on an octavo page with 9 point type size, his medium sentence occupies between one and two completed lines, and his long sentence more than two lines. He points out that people vary widely in their use of these three types of sentence.

3. The experiments

Such a spate of often contradictory recommendations as to the optimum length of sentence in written prose suggests that this is an important factor in 'good' writing, a factor that has been consistently neglected in second language teaching. It was decided, therefore, to set up experiments to discover

(a) what length of sentence a panel of experienced readers would prefer

(b) whether any significance could be attached to the positioning of short sentences in the text.

The following hypotheses would be tested :

Hypothesis 1 (Experiment 1)

Experienced readers will tend to reject the extremes (ie short and long sentences) in favour of passages with varied and medium-length sentences. There will be a preference for sentences of varied over those of medium length.

Hypothesis 2 (Experiment 2)

When considering passages with only sentences of varied length, experienced readers will tend to prefer those with short sentences at change of topic to those with short sentences randomly placed.

3.1 Method As the same base texts and subjects were used, the two experiments will be considered together.

3.1.1 Experiment 1 Fifty-four subjects were divided into three groups of 17, 18 and 19, and each group was presented with four versions of one of three different texts, the only variable being sentence length. Version A was written in sentences of varied length, Version B in medium-length sentences, Version C in long sentences and Version D in short sentences. The subjects were asked to rank the passage in order of preference, having been instructed to bear in mind their comprehensibility, readability, flow, clarity, and general 'pleasantness' or otherwise of style. The order of presentation was counter-balanced. The passages can be found in Appendix 2a.

3.1.2 Experiment 2 Fifty-five subjects in three groups of 21, 15 and 19 were presented in a similar way with two versions of one of the same three texts. Each version consisted of sentences of varied length, the only variable the positioning of the short sentences. The short sentences in Version A were placed at change of topic, while those in Version B were randomly placed. The subjects were given the same instructions as for Experiment 1, and the order of presentation was again counter-balanced. These passages can be found in Appendix 2b.

3.2 Subjects A group of subjects was needed who would be mother-tongue speakers of English, graduates, and used to the type of expository prose read by sixth-formers and

first year undergraduates. It was decided to ask for up to sixty volunteers from the staff of Moray House College of Education, Edinburgh. It was originally intended to use only teachers of Education, Biology and the Social Sciences, but this proved impossible owing to the difficulty of obtaining the volunteers, so it was necessary to use members of a variety of departments. Teachers of English and Linguistics, however, were not represented, as it was hoped that the subjects would approach the experiments with as few linguistic and stylistic preconceptions as possible. A fairly successful attempt was made to place an equal proportion of subjects from the different departments in each group.

3.3 Material Three short passages of approximately 200 to 250 words in length were chosen from secondary school textbooks and freely adapted for the purpose of the experiment. The passages were such that each consisted of at least three sub-divisions. Secondary school level texts were chosen to ensure that the subject matter was sufficiently simple for its presentation in short sentences to appear reasonably natural. From this point onwards the passages will be referred to as the earthquakes, soil and tornados passages.

3.3.1 Method of construction A major difficulty occurs with an experiment of this type : how can a syntactic variable (or variables) be manipulated without also changing the thematic meaning of the text (as defined by Leech, 1974)?

In an earlier pilot experiment, the simultaneous semantic structure representing deep level meaning was indicated on a network, from which the four different representations of the sequential structure at the surface level were derived. An attempt was made to control certain syntactic variables apart from sentence length, namely, type of clause and depth of embedding, but the result was that in some cases the order of presentation within the sentence differed between the passages, thus subtly altering the thematic meaning. It was therefore decided that, instead of using a network as a semantic base, the examples of Dautermann (1969) and Hunt (1971) would be followed.

Dautermann was aiming to determine the pattern of growth in syntactic maturity, so needed a method for enabling school-children to construct their own individual texts from a common semantic base. He used 76 'modified kernel sentences', to be combined by the subjects in their own way into a narrative. Hunt used a similar instrument to test the use of embedding transformations at different levels of maturity, this time using an expository rather than a narrative text.

In the present experiments, therefore, each passage was divided up into 32 or 33 modified kernels. Instead of requiring the subjects themselves to combine the kernels, as Dautermann and Hunt had done, the combinations were made by the experimenter manipulating the variable of sentence length. It should be emphasised, however, that these are definitely not kernel sentences in the transformational

sense : the division is purely pragmatic, each 'kernel' being a kind of minimal sentence, the internal constituents of which will not be radically re-positioned. The modified kernels for each passage can be seen in Appendix 2c.

Before constructing the texts from the kernels it was necessary to define the terms 'long', 'short', 'medium' and 'varied'. This was done arbitrarily, and is inevitably somewhat artificial, as in an authentic situation there is obviously no set point on one side of which a sentence is 'long' and on the other 'medium' or 'short'. The boundaries, however, were drawn up after considerable study of authentic text, and they correlate reasonably well with Wehrlich's (1976) definitions if it is calculated that a printed line contains on average about twelve individual words.

The experiment was concerned with short texts rather than individual sentences, so the criteria used are averages over the whole text, and are as follows :

In a text of approximately 200 to 250 words :

a) Short The sentences average less than ten words. No more than one sentence should exceed 15 words, but there is no lower limit. In no case should a sentence be three times or more the length of a juxtaposed sentence.

b) Medium The sentences average ten to 30 words. No sentence should be less than ten words and no sentence more than 30. In no case should a sentence of ten words juxtapose a sentence of 30 words.

c) Long The sentences average 30 words or more. Not more than one sentence should be less than 20 words, but there is no upper word limit. In no case should a sentence be three times or more the length of a juxtaposed sentence.

d) Varied The sentences average ten to 30 words. At least one sentence should be less than ten words, and at least one more than 30. In at least two places one sentence should juxtapose another three times or more its own length.

3.3.1.1 Difficulties specific to Experiment 1 No problems were experienced in producing the short sentence passages, as all that was necessary was to make each modified kernel a separate sentence. Similarly, the medium sentence passages presented no difficulty, as can be seen by the uniformity in sentence totals and average sentence lengths between the passages (tables 9, 10 and 11).

The long sentence passages were a different matter. Katz and Fodor (1963) suggested that every text could be reduced to a single sentence - indeed, that this possibility was the test as to whether a passage is a text or not - but they clearly could never have applied this to actual texts. Any attempt to do so would reveal that there are internal sub-divisions within all texts of any length across which it is impossible, or at least highly unacceptable, to connect sentences. Such boundaries are the boundaries of information blocks.

TABLE 9

PASSAGES FOR EXPERIMENT 1 : TOTAL WORDS

	VARIED	MEDIUM	LONG	SHORT
EARTHQUAKES	232	227	234	227
SOIL	250	244	252	248
TORNADOES	235	225	238	256

TABLE 10

PASSAGES FOR EXPERIMENT 1 : TOTAL SENTENCES

	VARIED	MEDIUM	LONG	SHORT
EARTHQUAKES	11	11	4	31
SOIL	12	12	5	31
TORNADOES	12	11	5	32

TABLE 11

PASSAGES FOR EXPERIMENT 1 : AVERAGE SENTENCE LENGTHS (IN WORDS)

	VARIED	MEDIUM	LONG	SHORT
EARTHQUAKES	21.09	20.63	58.50	7.32
SOIL	20.83	20.33	50.40	8.00
TORNADOES	19.58	20.45	47.60	8.00

Thus, with respect to these three passages, the information blocks within the earthquakes passage are sufficiently large for the text to be divided into only four sentences, with an average length of 58.50 words. The blocks within the other passages, however, are smaller, so that both need five sentences, with an average length of 50.4 words for the soil passage and 47.60 words for the tornadoes passage.

An examination of the long sentence version of the tornadoes passage (Appendix 2a) will clarify the problem. The first sentence is a general introduction to the subject of tornadoes, while the second and third sentences describe the appearance and explain the destructiveness of the tornado. A grammatical difference between the more general first sentence and the less general second and third sentences is that the first refers to 'tornadoes' and is plural in form, while the second and third refer to 'a tornado' and are singular. Although all three sentences are generic, there is no acceptable way in which the first and second sentences can be joined together.

The second and third sentences could possibly be conjoined, but the result would be unwieldy, and would contravene the condition (for the purposes of this experiment) on the long sentence text that no sentence should be three times or more the length of its immediate neighbour. The fourth sentence tells us where tornadoes occur, it being noteworthy that there is a change back to the plural 'tornadoes' here. The last sentence links two ideas, which,

according to one reading, could be considered coordinate : duststorms and dust devils. A coordinate relationship would allow conjoining, even though the resultant length would be somewhat unmanageable. An alternative reading would treat dust devils as subordinate to tornadoes, and dust storms as another form of atmospheric disturbance which, although probably less important than tornadoes, could not be considered as rhetorically subordinate. This type of ambiguity is common in poor writing, though it is obviously unfair to judge a passage which is just one small part of a larger text, and has in any case been somewhat freely adapted.

There seemed no way, therefore, of reducing the number of sentences in this passage, nor in the soil passage, so as to correspond more exactly to the earthquakes passage. The problem is especially marked in the tornadoes passage, where the average number of words per sentence is 10.09 fewer than in the earthquakes passage, but it was decided to use this difference to find out the subjects' reaction to relatively shorter 'long' sentences. It was predicted from preparatory pilot studies that they would be particularly sensitive to the lengths of such sentences. An addition can therefore be made to Hypothesis 1 :

Supplementary Hypothesis 1b

Concerning long sentence passages only, subjects will tend to prefer those with sentences towards the medium end of the cline to those towards the long end.

The question of the presence of lower level information blocks as referred to above will be taken up again in Chapter 9.

3.3.1.2 Difficulties specific to Experiment 2 Although an experienced writer could easily assign the short sentences to the most 'natural' positions by intuition alone, this will not suffice in experimental procedure. It was necessary, therefore, to use an explicit model to represent the places in the text where short sentences would indicate change of topic and where they would not. For this purpose an adaption of Grimes' Semantic Grammar of Propositions was chosen (See Grimes 1975). As this grammar will be the basis for the experiments of the next chapter, it will not be described in detail here. All that will be said about it is that its application to a text results in a hierarchical tree diagram consisting of layers of propositions which reveal the content structure of the passage. The tree diagram constructed for these three passages can be seen in Appendix 2d.

Inevitably, the application of such a model to a text is to a certain extent subjective, as an author seldom makes every connection explicit, especially at the higher levels of the tree. This may be especially true when using small extracts from a much longer text, in that connections which are obvious when viewing the text as a whole may not be so when viewing the extract in isolation. An example may be seen in the tornadoes passage. Here it is not

possible to determine exactly from the passage itself whether we are dealing with tornadoes, further sub-divided into three groups, ie tornadoes proper, and their related phenomena dust devils and duststorms, or with three separate phenomena related at a higher and more general level, that of atmospheric disturbances. Fortunately this problem does not affect the validity of the model for this particular experiment, but it should at least be borne in mind.

In Appendix 2d the numbers correspond to the modified kernels in Appendix 2c. The short sentences at change of topic were assigned to kernels at the highest level nodes, with the exception of the nodes representing the first sentences of the passages. These were omitted not because they were unsuitable places for short sentences, as they certainly are not, but because as each passage begins with a short introductory section which forms a separate information block, the first kernel could not be followed by a long enough sentence to qualify as 'short'. We are forced by the experimental procedure to adopt a rigid definition of the 'short' sentence : it should not be forgotten, however, that in reality length is represented as a cline, with a long transitional area between 'long' and 'short'.

The randomly placed short sentences were chosen from as low down on the tree as possible, though it was thought advisable not to place them at the bottom of a node just before a movement to a higher node, as this seemed intuitively to be a position where short sentences

would often be found. The only text which presented problems here was again the tornadoes passage. Because this is not a densely structured passage, having only six levels, it proved difficult to find suitable kernels to act as short sentences. Those finally chosen were not considered entirely satisfactory, as one (no.13) occurred just before an upward movement of two levels (the second largest in the passage), and they were not evenly distributed throughout the passage as in the other two passages. The final choice, however, was inevitable, for much the same reasons as those already given for the difficulty in producing suitable long sentences.

3.4 Results : statistical analyses

The raw results of Experiment 1 can be seen in Table 12 and those of Experiment 2 in Table 13.

TABLE 12

EXPERIMENT 1 : RAW RESULTS

(a) Earthquakes (total = 18)

	VERSIONS			
	A	B	C	D
1st	8	9	0	1
2nd	9	7	1	1
3rd	1	2	9	6
4th	0	0	8	10

(b) Soil (total = 19)

	VERSIONS			
	A	B	C	D
1st	9	8	2	0
2nd	8	7	3	1
3rd	1	4	10	4
4th	1	0	4	14

(c) Tornadoes (total = 17)

	VERSIONS			
	A	B	C	D
1st	4	10	3	0
2nd	6	5	6	0
3rd	6	2	4	5
4th	1	0	4	12

(d) Combined (total = 54)

	VERSIONS			
	A	B	C	D
1st	21	27	5	1
2nd	23	19	10	2
3rd	8	8	23	15
4th	2	0	16	36

KEY

A = Varied
B = Medium
C = Long
D = Short

TABLE 13

EXPERIMENT 2 : RAW RESULTS

	1	2	3	4
CHANGE OF TOPIC	17	12	14	43
RANDOM	4	3	5	12

KEY

1 Earthquakes : Total = 21
2 Soils : Total = 15
3 Tornadoes : Total = 19
4 Combined : Total = 55

3.4.1 Experiment 1 The following tests were used to test the significance of the results for Experiment 1. In all cases a probability level of 0.05 is required to indicate significance, while a level of 0.01 would indicate high significance.

3.4.1.1 Kendall's Coefficient of Concordance W (See Appendix 3a). The application of this test indicates a strong degree of association between the judgments made on each passage. The highest degree of association is found with the earthquake passage (W = 0.63) and the lowest with the tornadoes passage (W = 0.49) with the soils passage (W = 0.54) between the two. Thus all three are well below the 0.01 level of significance. It therefore seems legitimate to combine them, with the result again that the association is strongly marked (W = 0.51) and highly significant ($p = <0.01$). The agreement of the judges is thus shown to be considerably higher than would occur by chance alone. It should be noted, however, that this cannot be interpreted as meaning that the ordering they have decided on is in any way correct, but only that they are applying essentially the same standard in their judgements.

3.4.1.2 Chi-square applied to each passage (See Table 14)
Null Hypothesis (H_o) : Taking each passage separately, the proportion of first, second, third and fourth choices made for each of the four sentence lengths is the same in all cases.

Alternative Hypothesis (H_1) : Taking each passage separately, the proportion of first, second, third and fourth choices made for each of the four sentence lengths differs from case to case.

As expected frequencies in no case exceeded 4.5, it was decided to combine cells in columns 1 and 2, and those in columns 3 and 4. (first and second choices, third and fourth choices). The combined cells produced expected frequencies of above 8.50, well beyond the critical level of 5.00.

TABLE 14

THE CHI-SQUARE TEST APPLIED TO EACH PASSAGE

(a) Earthquakes

	1st	2nd	3rd	4th	
A	4.5 / 8	4.5 / 9	4.5 / 1	4.5 / 0	18
B	4.5 / 9	4.5 / 7	4.5 / 2	4.5 / 0	18
C	4.5 / 0	4.5 / 1	4.5 / 9	4.5 / 8	18
D	4.5 / 1	4.5 / 1	4.5 / 6	4.5 / 10	18
	18	18	18	18	72

	1st & 2nd	3rd & 4th	
A	9.0 / 17	9.0 / 1	18
B	9.0 / 16	9.0 / 2	18
C	9.0 / 1	9.0 / 17	18
D	9.0 / 2	9.0 / 16	18
	36	36	72

(b) Soil

	1st	2nd	3rd	4th	
A	4.75 / 9	4.75 / 8	4.75 / 1	4.75 / 1	19
B	4.75 / 8	4.75 / 7	4.75 / 4	4.75 / 0	19
C	4.75 / 2	4.75 / 3	4.75 / 10	4.75 / 4	19
D	4.75 / 0	4.75 / 1	4.75 / 4	4.75 / 14	19
	19	19	19	19	76

	1st & 2nd	3rd & 4th	
A	9.5 / 17	9.5 / 2	19
B	9.5 / 15	9.5 / 4	19
C	9.5 / 5	9.5 / 14	19
D	9.5 / 1	9.5 / 18	19
	38	38	76

(c) Tornadoes

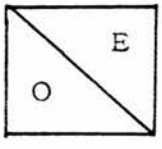
	1st	2nd	3rd	4th	
A	4.25 / 4	4.25 / 6	4.25 / 6	4.25 / 1	17
B	4.25 / 10	4.25 / 5	4.25 / 2	4.25 / 0	17
C	4.25 / 3	4.25 / 6	4.25 / 4	4.25 / 4	17
D	4.25 / 0	4.25 / 0	4.25 / 5	4.25 / 12	17
	17	17	17	17	68

	1st & 2nd	3rd & 4th	
A	8.5 / 10	8.5 / 7	17
B	8.5 / 15	8.5 / 2	17
C	8.5 / 9	8.5 / 8	17
D	8.5 / 0	8.5 / 17	17
	34	34	68

(d) Combined

	1st	2nd	3rd	4th	
A	13.5 / 21	13.5 / 23	13.5 / 8	13.5 / 2	54
B	13.5 / 27	13.5 / 19	13.5 / 8	13.5 / 0	54
C	13.5 / 5	13.5 / 10	13.5 / 23	13.5 / 16	54
D	13.5 / 1	13.5 / 2	13.5 / 15	13.5 / 36	54
	54	54	54	54	216

KEY



O = observed figure

E = expected figure

Results show a difference between the passages - for 'earthquakes' X^2 = 50.20, for 'soils' X^2 = 37.68 and for 'tornadoes' X^2 = 27.52 (all with 3df). All these values of X^2, however, are significant below the 0.001 level (X^2 = 16.27), so it was considered legitimate to combine the figures. The result of this, X^2 = 126.83 at 9df, also shows a highly significant result well below the 0.001 level. The Null Hypothesis can therefore be rejected, showing that the choices made are strongly dependent on the sentence lengths being judged.

3.4.1.3 Chi-square applied to each sentence length (See Table 15)

Null Hypothesis (H_o) : Taking each sentence length separately, the proportion of first, second, third and fourth choices made in each passage is the same in all cases.

Alternative Hypothesis (H_1) : Taking each sentence length separately, the proportion of first, second, third and fourth choices made in each passage differs from case to case.

As 50% and more of the cells in each case had expected frequencies of less than 5.00, it was decided to combine cells. The most suitable combinations proved to be third and fourth choice for varied and medium sentences and first and second choice for long and short sentences. Even after these combinations, however, the proportion of cells with an expected frequency of less than five was still high. Only the long sentence matrix showed the recommended figure of below 20% (11.11%), with medium and mixed at 33.33% and short still as high as 55%. Everett (1977 : 4), however, suggests that this may not be as important for the reliability of the test as was previously supposed. He points out that '... recent work by Lewontin and Felsenstein (1965), Slakter (1966) and others shows that many of the expected values may be as low as unity without affecting the test greatly'. It was therefore decided to carry out the tests, with the proviso that except for the long sentence version they may not be ideally reliable.

TABLE 15

THE CHI-SQUARE TEST APPLIED TO EACH SENTENCE LENGTH

(a) Varied (A)

	1st	2nd	3rd	4th	
EARTHQUAKES	7.0 / 8	7.67 / 9	2.67 / 1	0.67 / 0	18
SOIL	7.39 / 9	8.09 / 8	2.81 / 1	0.70 / 1	19
TORNADOES	6.61 / 4	7.24 / 6	2.52 / 6	0.63 / 1	17
	21	23	8	2	54

	1st	2nd	3rd & 4th	
EARTHQUAKES	7.00 / 8	7.67 / 9	3.33 / 1	18
SOIL	7.39 / 9	8.09 / 8	3.52 / 2	19
TORNADOES	6.61 / 4	7.24 / 6	3.15 / 7	17
	21	23	10	54

(b) Medium (B)

	1st	2nd	3rd	4th	
EARTHQUAKES	9.0 / 9	6.33 / 7	2.67 / 2	0 / 0	18
SOIL	9.5 / 8	6.69 / 7	2.81 / 4	0 / 0	19
TORNADOES	8.5 / 10	5.98 / 5	2.52 / 2	0 / 0	17
	27	19	8	0	54

	1st	2nd	3rd & 4th	
EARTHQUAKES	9.0 / 9	6.33 / 7	2.67 / 2	18
SOIL	9.50 / 8	6.69 / 7	2.81 / 4	19
TORNADOES	8.50 / 10	5.98 / 5	2.52 / 2	17
	27	19	8	54

(c) Long (C)

	1st	2nd	3rd	4th	
EARTHQUAKES	1.67 / 0	3.33 / 1	7.67 / 9	5.33 / 8	18
SOIL	1.76 / 2	3.52 / 3	8.09 / 10	5.63 / 4	19
TORNADOES	1.57 / 3	3.15 / 6	7.24 / 4	5.04 / 4	17
	5	10	23	16	54

	1st & 2nd	3rd	4th	
EARTHQUAKES	5.00 / 1	7.67 / 9	5.33 / 8	18
SOIL	5.28 / 5	8.09 / 10	5.63 / 4	19
TORNADOES	4.72 / 9	7.24 / 4	5.04 / 4	17
	15	23	16	54

(d) Short (D)

	1st	2nd	3rd	4th	
EARTHQUAKES	0.33 / 1	0.67 / 1	5.00 / 6	12.00 / 10	18
SOIL	0.35 / 0	0.70 / 1	5.28 / 4	12.67 / 14	19
TORNADOES	0.32 / 0	0.63 / 0	4.72 / 5	11.33 / 12	17
	1	2	15	36	54

	1st & 2nd	3rd	4th	
EARTHQUAKES	1.00 / 2	5.00 / 6	12.00 / 10	18
SOIL	1.06 / 1	5.28 / 4	12.67 / 14	19
TORNADOES	0.94 / 0	4.72 / 5	11.33 / 12	17
	3	15	36	54

KEY

E
O

O = observed figure

E = expected figure

Results show a strong similarity between the three passages when medium and short sentence choices are made (medium : X^2 = 1.53 ; short : X^2 = 2.98), and a greater difference when varied and long choices are made (varied : X^2 = 8.95 ; long : X^2 = 11.25) (all with 4 df). As the critical X^2 value for a significance value of 0.05 is 9.49, the Null Hypothesis therefore stands for medium, short and varied sentences, but must be rejected for long sentences. This means that there is no significant difference from passage to passage between the positions given by the judges to medium, mixed and short sentences, but that in the positions given to long sentences there is a difference between the passages which is greater than would be likely by chance alone.

3.4.1.4 Analysis of residuals (See Table 16, Appendix 3) This method has been used to isolate the sources of association in the long and mixed passages (See Everett 1977 : 47). If the variables are independent, the adjusted residuals should be approximately normally distributed, with mean 0 and standard deviation 1, so that any cell with a considerably higher or lower value will be significant. Everett suggests the 5% standard normal deviate, ie 1.96, as the limit to significance.

TABLE 16

ADJUSTED RESIDUALS

(a) Long (C)

	1st & 2nd	3rd	4th
EARTHQUAKES	**-2.59**	0.77	1.68
SOIL	-0.17	1.10	-1.01
TORNADOES	**2.77**	-1.94	-0.67

(b) Varied (A)

	1st	2nd	3rd & 4th
EARTHQUAKES	0.59	0.77	-1.75
SOIL	0.94	-0.05	-1.11
TORNADOES	-1.57	-0.74	**2.91**

KEY

[ ] Significant value

Applying the analysis to the long sentence passage (Table 16a), it can be seen that two cells are particularly significant - the earthquakes combined first and second choices, with an unusually low number, and the tornadoes combined first and second, with an unusually high number. Thus considerably more judges made the long sentence version their first choice in the tornadoes passage than in the other two passages, and rather fewer chose the long sentence version in the earthquakes passage.

Although no statistically significant difference from passage to passage was indicated for varied passages, of the remaining three, these came closest to significance, so it is useful to apply the analysis of residuals to these as well (Table 16b). Again it is the tornadoes passage which contributes most to any divergency, there being a significantly high number for the combined third and fourth choices. This correlates well with the analysis for long sentences. When considering the tornadoes passage, more of the judges tended to place the long sentence versions in first and second position, and more placed the varied sentence versions in third and fourth position.

3.4.1.5 The Kolmogorov-Smirnov One-Sample Test (Goodness of fit) (See Table 17) Unlike the Chi-Square Test, this test avoids the necessity of combining cells and thus losing information.

Null Hypothesis (H_o) : Taking each version of each passage separately, there is no difference in the expected number of choices made for each version, and any observed differences are merely chance variations to be expected in a random sample where $f_1 = f_2 = f_3 = f_4$

Alternative Hypothesis (H_1) : Taking each version of each passage separately, the frequencies f_1, f_2, f_3, f_4 are not all equal.

TABLE 17

THE KOLMOGOROV-SMIRNOV ONE-SAMPLE TEST

		A	B	C	D	
EARTHQUAKES	D	0.44	0.39	0.44	0.39	Where D = 0.31, p = 0.05
N = 18	P	$<$.01	$<$.01	$<$.01	$<$.01	D = 0.37, p = 0.01
SOIL	D	0.39	0.29	0.24	0.49	Where D = 0.30, p = 0.05
N = 19	P	$<$.01	NOT SIGN	NOT SIGN	$<$.01	D = 0.36, p = 0.01
TORNADOES	D	0.19	0.38	0.07	0.50	Where D = 0.32, p = 0.05
N = 17	P	NOT SIGN	.01	NOT SIGN	$<$.01	D = 0.38, p = 0.01
COMBINED	D	0.31	0.35	0.22	0.44	Where D = 0.19, p = 0.05
N = 54	P	$<$.01	$<$.01	.01	$<$.01	D = 0.22, p = 0.01

D = maximum deviation of O from E

Results show that

(a) For <u>varied sentences</u> (A) the Null Hypothesis can be rejected on two passages (earthquakes and soils) and upheld on one (tornadoes). In the cases where it is rejected, p $=<$0.01.

(b) For <u>medium sentences</u> (B) the Null Hypothesis can be rejected on two passages (earthquakes and tornadoes) and upheld on one (soils), the latter case, however, falling only fractionally above a significance level of 0.05. Where H_o is rejected, p $=<$0.01.

(c) For <u>long sentences</u>, the Null Hypothesis can be rejected only on the earthquakes passage (p $=<$0.01). On the soils and tornadoes passages it is upheld.

(d) For <u>short sentences</u>, the Null Hypothesis can be rejected on all passages, with p $=<$0.01.

When all three passages are combined, the Null Hypothesis can in all cases be rejected, with p at 0.01 or below.

3.4.1.6 Somers's 'd' (See Table 18, Appendix 3d). This analysis gives a measure of association between the two variables of sentence length and order of choice. This particular measure was chosen because of its suitability for situations where there are two variables, one explanatory (x) and the other dependent (y) (Everett, 1977 : 63). For the purpose of the analysis, the x variable of sentence length has been re-ordered so as to correspond to the frequency of first choice position when all three passages are combined (ie medium, varied, long, short).

When the data is concentrated in the upper-left to lower-right diagonal d has the value of +1, and when the two variables are completely independent it has zero value. Thus all three versions show a strong correlation between the ranking choice and the sentence length. Earthquakes = 0.61, soil = 0.59, and tornadoes = 0.59 ; all three passages combined = 0.60.

TABLE 18

EXPERIMENT 1 : SOMERS'S 'd'

(a) Earthquakes

y \ x	B	A	C	D
1st	9	8	0	1
2nd	7	9	1	1
3rd	2	1	9	6
4th	0	0	8	10

d = 0.61

(b) Soil

y \ x	B	A	C	D
1st	8	9	2	0
2nd	7	8	3	1
3rd	4	1	10	4
4th	0	1	4	14

d = 0.58

(c) Tornadoes

y \ x	B	A	C	D
1st	10	4	3	0
2nd	5	6	6	0
3rd	2	6	4	5
4th	0	1	4	12

d = 0.59

(d) Combined

y \ x	B	A	C	D
1st	27	21	5	1
2nd	19	23	10	2
3rd	8	8	23	15
4th	0	2	16	36

d = 0.60

3.4.1.7 Binomial Test applied to varied and medium sentences (See Table 19). The raw results suggest no clear difference in the choices between the varied and medium sentences, so the Binomial Test was applied to find out whether any significant differences could be detected. The medium and varied choices were therefore extracted from the total choices and compared with each other.

TABLE 19

THE RELATIVE POSITIONS OF VARIED AND MEDIUM SENTENCES

	A	B	TOTAL
EARTHQUAKES	9	9	18
SOIL	11	8	19
TORNADOES	4	13	17
COMBINED	24	30	54

KEY

A = Varied sentences placed before medium sentences

B = Medium sentences placed before varied sentences

Null Hypothesis (H_o) : $p_1 = p_2 = \frac{1}{2}$ ie there is no difference between the probability of choosing the medium sentence versions (p_1) and the probability of choosing the varied sentence versions (p_2)

Alternative Hypothesis (H_1) : $p_1 > p_2$, or $p_2 > p_1$, ie the probability of choosing the medium sentence versions is greater than the probability of choosing the varied sentence versions, or vice versa.

As no prediction was being made as to which version would be preferred, a two-tailed test was used, and the resultant probabilities were therefore doubled. Results show that the Null Hypothesis is upheld in all cases, though the probability level for the tornadoes passage is only barely above significance ($p = 0.06$). This indicates that the judges as a whole tend not to prefer medium over varied sentences, or varied over medium sentences, though there is a slight though insignificant tendency towards choosing the medium version in the tornadoes passages.

3.4.2 Experiment 2 The following tests were used to test the significance of the results of Experiment 2. In all cases a probability level of 0.05 is required to indicate significance, while a level of 0.01 would indicate high significance.

3.4.2.1 Chi-Square Test (See Table 20)

Null Hypothesis (H_o) : The proportion of choices made for short sentences at change of topic and short sentences randomly placed is the same from passage to passage.

Alternative Hypothesis (H_1) : The proportion of choices made for short sentences at change of topic and short sentences randomly placed differs from passage to passage.

Although three of the six cells contain expected values below 5.00, it is not possible in this case to combine cells. However, according to Everett (1977 : 40), Lewontin and Felsenstein (1965) have stated : "The 2xc tables can be tested by the conventional chi-square criterion if all the expectations are one or greater", and it was therefore decided that the test was valid, especially as here no expected value falls less than 1.73 below the recommended value.

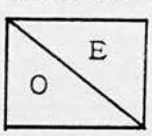
TABLE 20

CHI-SQUARE TEST : THE POSITIONING OF SHORT SENTENCES

	1	2	3	
AT CHANGE OF TOPIC	16.42 / 17	11.73 / 12	14.85 / 14	43
RANDOM	4.58 / 4	3.27 / 3	4.15 / 5	12
	21	15	19	55

KEY

1 Earthquakes
2 Soil
3 Tornadoes



O = observed figure
E = expected figure

Results show that, with 2df, p = 0.85, and therefore the Null Hypothesis is upheld. This means that there is no significant difference from passage to passage between the choices made by the judges.

3.4.2.2 The Binomial Test (See Table 13)

Null Hypothesis (H_o) : $p_1 = p_2 = \frac{1}{2}$ ie there is no difference between the probability of choosing the version with short sentences at change of topic (p_1) and the probability of choosing that with short sentences randomly placed (p_2).

Alternative Hypothesis (H_1) : $p_1 > p_2$, ie the probability of choosing the version with the short sentences at change of topic is greater than the probability of choosing that with short sentences randomly placed.

Results show that, with a significance level of 0.05, the Null Hypothesis can be rejected in all cases. The soils and tornadoes passages, with a probability of 0.020 and 0.034 respectively, have a significance level of below 0.05, while the earthquakes passage shows a level of 0.004, a level well below 0.01. When the three passages are combined, the result is highly significant ($p = < 0.00003$), showing that as far as these passages are concerned the judges overwhelmingly prefer the versions with short sentences at change of topic.

4. General discussion of results An attempt will be made briefly to summarise the results of the above experiments, according to the statistical tests.

4.1 Experiment 1 : Hypothesis 1a The first part of this hypothesis, ie that experienced readers will tend to reject the extreme, is upheld. On all three passages the short sentence versions were overwhelmingly rejected, and on two out of three of the passages the long sentence versions were rejected also. The 'Somers's d'analysis illustrates how the short sentence versions have a strong tendency towards being placed in fourth and final position, while the long sentence versions have a tendency less strong but still pronounced towards being placed in third position.

Three of the judges made their decisions explicit in the following written comments :

> I dislike <u>equally</u> too many short sentences and the reverse.
>
> I dislike the lack of flow from overuse of simple sentences.
>
> It was difficult to separate numbers three and four in preference. Both bore, one by over-short sentences the other by over-long.

Most readability formulae, however, recommend short sentences, and when two well-known formulae were applied to the different versions of these passages (Table 21) the short sentence version appeared as without doubt the easiest to read. It seems questionable, however, whether sentence length and word length criteria are sufficient. It is true that short sentences usually imply simple syntax, but other considerations are also important for ease of reading, eg interest, and the smooth connection of ideas. These influences require psycho-linguistic rather than purely linguistic considerations.

The second part of the hypothesis, ie that experienced readers will prefer sentences of varied to those of medium length, has to be rejected. Although there is a predominance of varied and/or medium sentences in first and second positions, a finding illustrated by the 'Somers's d'Analysis, on the application of a two-tailed Binomial Test no overall significant difference between the two was evident, either for individual passages or for all three combined. However, the tornadoes passage in isolation does stand out as being somewhat different, in that there is a relatively small

number of first choices for the varied version, and a relatively large number for the medium version. Although without the tornadoes passage there is a slight preference for the varied sentence, this is not statistically significant, and it is unlikely that it would have been even if the tornadoes passage had conformed more closely to the other two.

TABLE 21

THE READABILITY OF THE PASSAGES, AS MEASURED BY TWO READABILITY FORMULAE

(a) Gunning Fog Index

	A	B	C	D
EARTHQUAKES	13.62	13.36	28.69	7.86
SOIL	12.06	12.22	23.80	6.74
TORNADOES	13.62	13.90	23.91	7.72

(b) Flesch R.E. Formula

	A	B	C	D
EARTHQUAKES	48.83	48.39	10.81	62.64
SOIL	61.04	60.35	32.14	74.89
TORNADOES	55.49	56.00	28.07	70.83

Whether the experiment adequately reflects reality in this respect, however, is another matter. A major problem is that of the length of the passage. It could

be that in a text of normal length, eg an article, a chapter or a book, medium sentences would become monotonous, and the reader would then more definitely prefer a variety of sentence length. Unfortunately this is impossible to test with the method that has been used above. Even these passages proved rather too long for some of the judges to handle easily, and several complained of the boredom involved in reading and judging four passages with the same content.

4.2 Hypothesis 1b This hypothesis is upheld. The Chi-Square Test on each sentence length shows that only the long sentence versions of the passages varied significantly from each other in the judgments. The tornadoes passage, with the shortest 'long' sentences (47.60 average word length) showed the greatest number of long sentence choices in first and second position (nine out of the 17 total choices), while the earthquakes passage with the longest 'long' sentences (58.50 average word length) showed the smallest number (one out of 18 total choices). The analysis of residuals reinforces the picture, representing the earthquakes and tornadoes passages as mirror-images of each other. Moreover, the Kolmogorov-Smirnov Test indicates that the long sentence judgments become more variable as the sentence lengths decrease - whereas in the earthquakes passage the long sentence versions are solidly rejected, in the tornadoes passage the choices spread out, the observed figures becoming very close to the expected figures.

It is interesting to note that as this happens it is the varied sentence version that is displaced in the

tornadoes passage from its favoured position in first and second place, rather than the medium sentence version. This may be because as the sentences of the long sentence version become shorter, they become less easy to differentiate from the medium sentences, but it is more likely to be something intrinsic to the tornadoes passage itself.

4.3 Experiment 2 : Hypothesis 2 This hypothesis is upheld. The Binomial Test shows that a highly significant majority of judges preferred the versions with short sentences at change of topic to those with short sentences randomly placed.

Although the results for all three passages are separately significant, there is nevertheless a difference between them, the probability ranging from one of 0.004 for the earthquakes passage to one as high as 0.034 for the tornadoes passage. The reason for this almost certainly lies with the differing structures of the passages, already discussed in some detail. No entirely satisfactory version for the short sentences randomly placed could be found for the tornadoes passage, but in spite of this the judges still rejected it decisively. This strongly suggests that the short sentence can have as one of its functions the marking of change of topic, and if used in this way comprehension will be facilitated. Indeed, the short sentence as defined here will be considered as a marker of change of topic in the work to be pursued in the next chapter.

CHAPTER 6

THE INFORMATION BLOCK

1. Introduction

Theoretical linguistics has traditionally concerned itself with the sentence, while rhetoric has shown more interest in the paragraph and the development of the text as a whole. Little or no recognition has been given to the chunking of text, either between the sentence and the paragraph, above the paragraph, or independently of it. In this chapter an attempt will be made to throw some light on the structure of expository prose principally by concentrating on boundaries : the boundaries between chunks of prose or information blocks.

1.1 The information block and the paragraph This chapter is not primarily concerned with the paragraph, but as human subjects are involved the use of the term cannot be avoided. The concept of the paragraph exists in the mind of every educated mother-tongue speaker who puts pen to paper, and in order to pursuade such a subject to divide the text up into its constituent chunks it is necessary to talk in terms of paragraphs. It should be stressed again, however, that this writer sees the paragraph as no more than an orthographic convention, its main purpose being to indicate hierarchical organisation in the text and thus to aid the reader in his interpretation of it. Of course, the paragraph may have

other purposes, such as emphasis, or the marking of direct speech, but we are concerned here with expository prose only, for which this definition is in general valid.

A paragraph of normal length can usually be further sub-divided into lower level chunks, which have been termed variously 'sub-paragraphs' (by teachers of writing), 'discourse blocs' (Pitkins), 'discourse units' (Kaplan), 'clusters' (Jones and Faulkner), etc. It is being suggested here, however, that these lower level divisions are essentially the same as 'paragraph' divisions, the significant point being that some divisions are chosen by the writer as paragraph boundaries and some are not. What is, therefore, being investigated here is the chunking of text into information blocks of various sizes. The investigation is pursued mainly by examining the transitions between individual blocks. Some criteria for putting certain minor chunks together into major chunks, which are then orthographically marked off as paragraphs, will be examined later, in Chapter 8.

1.2 The investigation For the purposes of the investigation three different passages have been analysed. Each is a passage of authentic and unmodified expository prose extracted from a text of a type likely to be required reading for students of Geography in the Sixth Form or the first year of an undergraduate course. Each is between 690 and 820 words long, and is so constructed that it can be divided into five or six average length paragraphs. All three

passages can be seen in Appendix 4.

The investigation follows three parallel paths. Firstly, the passages are divided into chunks according to formal linguistic signals which can be identified in the text itself. Secondly, each passage is converted into a tree diagram based on Grimes' Semantic Grammar of Propositions (Grimes 1975), so that the passages can be divided into chunks according to upward movements in the tree structure. Lastly, an experiment is performed on each passage, in which subjects are asked to divide it hierarchically, at up to four levels, first into paragraphs, and then into sub-paragraph chunks. Taking native-speaker judgment (where sufficient agreement is achieved) as authoritative, relationships are then identified between the three analyses in an attempt to find out what the author's chunking consists of and how he makes those chunks evident to the reader.

From now on, these three passages will be referred to as Passage A (Soil Erosion), Passage B (Definition of Urban Areas) and Passage C (The Forest Community).

2. The Analysis according to signals identified in the text

When a reader interprets an expository text, he consciously or unconsciously uses signals, both implicit and explicit, as guides to the organisation of the information contained within it, and if these signals are not clear to him he will have difficulty in following the description, analysis or argument presented by the author. A well-written passage,

therefore, should present its material in clearly organised and signalled chunks, so that at no point does its reader become lost or confused.

An attempt is made here to divide up each of the three passages into chunks, on the basis of formal linguistic markers observable in the text itself. The assumption is being made that if we can identify some of the most effective signals of chunking, students can then be taught to reproduce these signals, and thus improve their own writing skills. Effectively placed signals, of course, are no guarantee that the chunks themselves will be well organised, but they are at the very least an aid to the reader, especially when the writer's skills are still not fully developed.

As this work is aiming at a practically biased end-product, the classification of signals chosen is based partly on their ease of manipulation by the writer. The major division is between intrinsic markers of content, which can be identified linguistically but are nevertheless part of the semantic content of the text, and inseparable from it, and extrinsic markers of cohesion, which are more independent of the semantic content of the text and are often manipulated consciously by the writer. It is not suggested, of course, that the list is final - only those signals were chosen that seemed most common, and appeared several times in the sample passages. Important areas have been ignored, eg the distribution of given and new information and structural parallelism, but a study of this kind inevitably entails selection.

The classification adopted is as follows :

A Extrinsic signals

(i) The author's own paragraphing

(ii) Short sentences

(iii) Contrast words

(iv) Text reference signals

a) Cataphoric

b) Anaphoric

c) Two-way

B Intrinsic signals

(i) Verb tense change

(ii) Beginning and end of lexical cohesion chains

(iii) Movement from particular to general

In the case of some of the above sub-divisions, the texts were divided for a more convenient analysis into T-units (minimal terminal units) rather than sentences. The minimal terminal unit was a measure first used by Kellogg Hunt in his analysis of syntactic maturity in school children. He defines it as '. . . exactly one main clause plus whatever subordinate clauses are attached to that main clause' (Hunt 1970 : 197). When a T-unit number is being referred to it will be preceded by the initials TU, and when a sentence number is being referred to it will be preceded by the initial S. The division of the three passages into T-units can be seen in Appendix 4b.

2.1 Extrinsic signalling Extrinsic signalling presents fewer problems of identification than intrinsic signals, so will be dealt with first. Reference can be made to Table 22 for a complete break-down of all signals.

2.1.1 The author's own paragraphing As all three passages were chosen because they were judged to be 'well-written', it must be assumed that, for the writer at least, the paragraphing is 'correct', and serves a definite function in interpretation. Nevertheless, as will be seen later, some of the paragraph divisions made by the authors were not chosen by a majority of the subjects in Experiment 3. Examples of these are the paragraph beginning at S18 in Passage A, and in Passage B that beginning at S35. Both are probably examples of paragraphs marking emphasis, a type not fully considered here.

However, it should not be thought that in all cases the author's own paragraphing is inevitably correct, as paragraphing in written text does not seem to be part of a native speaker's competence in the way that ability to manipulate grammatical structure is. Paragraphing in the sense that we are using it here is, like writing itself, a skill learned after the mother tongue has been acquired but as, like writing, it is related to speech, it is based to a certain extent on innate competence (in the Chomskeyan sense), but also on cognitive ability. Thus, as pointed out by Christensen, writers do at times paragraph inappropriately, and it should therefore not be assumed that it is possible to build up a theory of paragraphing simply on the basis of the written production of native speakers.

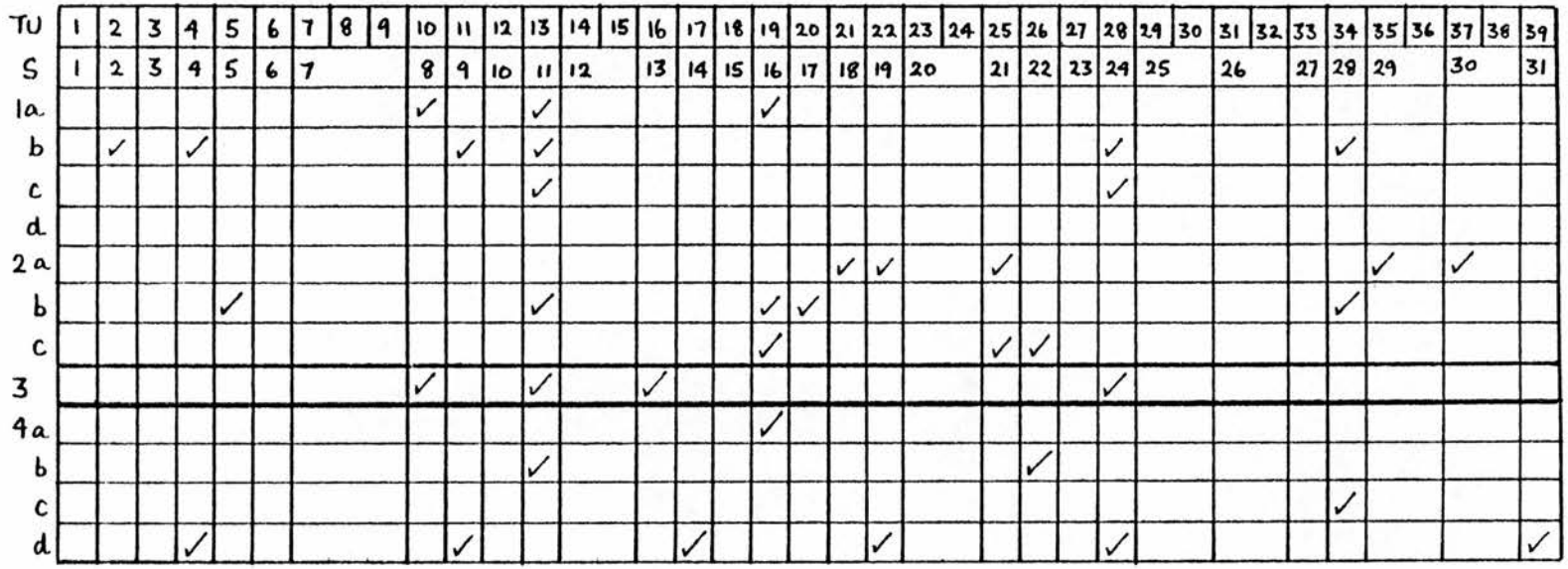
TABLE 22 POSITION OF SIGNALS AND BOUNDARY CHOICES

TU	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
S	1	2	3		4		5	6	7	8	9	10	11	12	13	14	15	16			17			18	19			20	21	22	23		24	25		26	27	28		29	30		31		32	33		34	35		36
1a																								✓				✓									✓															
b										✓						✓													✓				✓				✓															
c											✓																						✓																		
d																✓								✓				✓																									
2a											✓	✓																			✓			✓							✓											
b		✓									✓			✓				✓		✓								✓														✓									✓		✓
c											✓			✓		✓							✓																												✓
3														✓										✓				✓										✓													
4a																												✓																						
b										✓	✓																											✓												
c																								✓																											
d														✓		✓			✓		✓																													✓					

(a) PASSAGE A

TU	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
S	1	2	3	4	5	6		7	8	9	10		11	12	13	14		15	16	17	18	19	20		21		22	23	24	25		26	27	28	29	30	31	32	33	34	35	36	37	38	39
1a													✓							✓							✓											✓		✓					
b															✓																								✓						
c															✓																								✓						
d	✓												✓										✓				✓					✓			✓			✓				✓	✓		✓
2a			✓																				✓		✓			✓	✓								✓			✓				
b			✓							✓										✓						✓									✓			✓				✓			
c			✓					✓					✓							✓				✓														✓						
3													✓							✓							✓										✓						
4a																											✓																		
b													✓							✓									✓								✓	✓						
c																																														
d								✓	✓				✓																	✓		✓			✓					✓	✓		✓

(b) PASSAGE B



(c) PASSAGE C

KEY

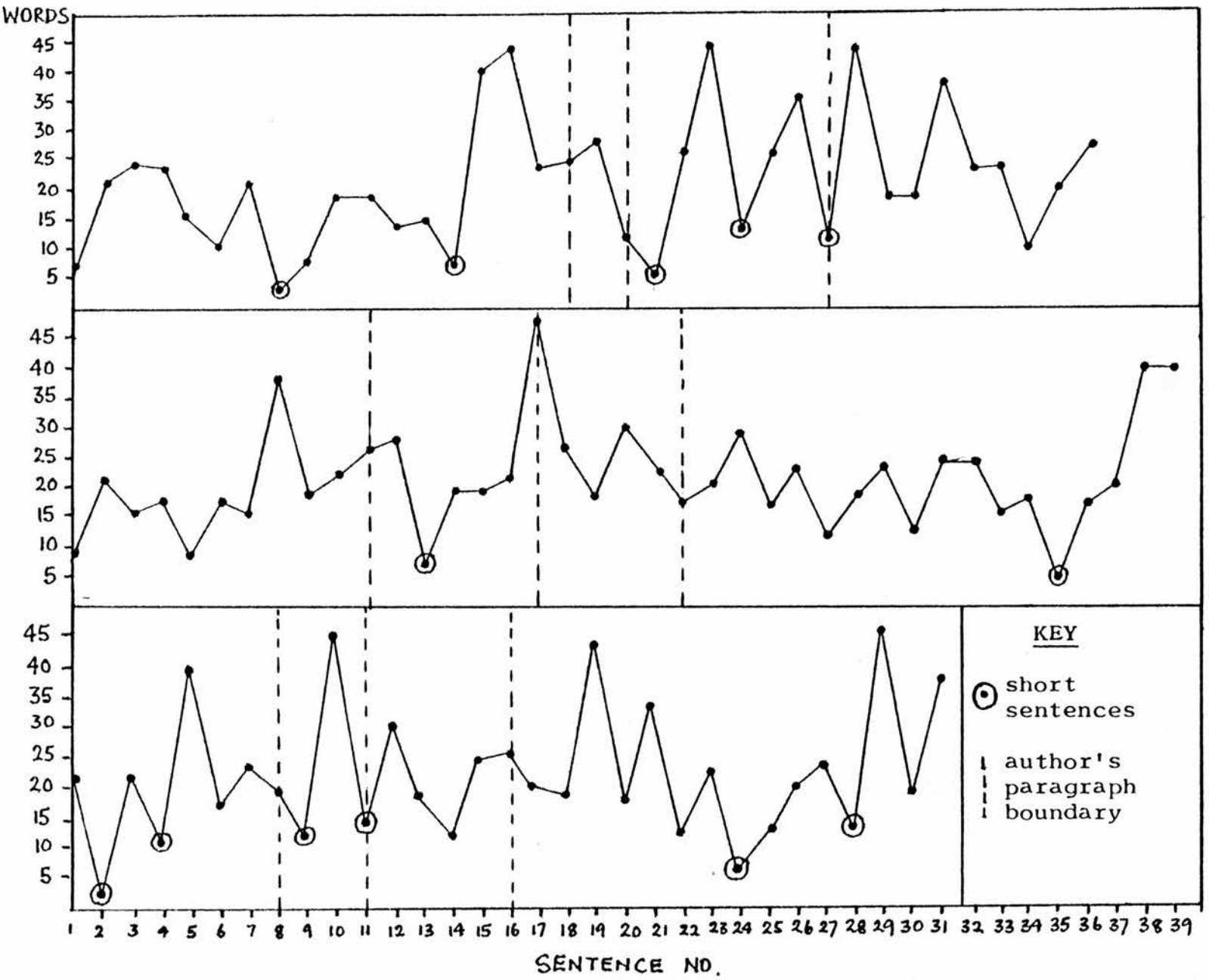
TU = T-Unit
S = Sentence
1a = Author's paragraphing
1b = Short sentences
1c = Contrast words
1d = Cataphoric reference signals
2a = Verb tense changes
2b = Lexical cohesion chains
2c = Particular/general
3 = Grimes Index of 0.16+
4a = Subjects' paragraphing: 1st Division } 50%+
4b = Subjects' paragraphing: 2nd Division }
4c = Subjects' paragraphing: 3rd Division } 40%+
4d = Subjects' paragraphing: 4th Division }

2.1.2 Short sentences (Figure 20) Following on from Experiment 2, short sentences are being considered here as markers of chunking. The same criteria for a short sentence are being used as before, namely that it should be one-third or less the length of the sentence on either one or both sides of it.

This is a marker which is used to a much greater extent by some writers than by others. All three writers here use it, but it is commonest in Passage C, with six examples, followed by Passage A (five examples) and then Passage B (two examples). When one considers that this is in inverse relationship to the total number of words, the difference between the passages is in fact greater than at first appears.

A short sentence is not necessarily used by a writer as the first sentence of a paragraph - indeed, it is more likely to appear contrapuntally, as it were, to point out change of topic within rather than between paragraphs. It is used most markedly in this way by the writer of Passage C, of whose six short sentences, all of which clearly signal transitions, only one begins a paragraph. Perhaps the short sentence is used so much by this writer because he has not made use of other signals to the same extent as the other writers have done, text reference signals in particular being almost completely absent.

FIGURE 20 SHORT SENTENCES IN THE THREE PASSAGES



PASSAGE A

Total words = 773
Av.sentence length = 21.47
Av.T.U. length = 15.46

PASSAGE B

Total words = 811
Av.sentence length = 20.79
Av.T.U. length = 18.02

PASSAGE C

Total words = 691
Av.sentence length = 22.29
Av.T.U. length = 17.72

In Passage A, too, short sentences tend to be in the middle rather than at the beginning of the author's paragraphs, with only one out of five appearing in first position. For this writer, also, the short sentence is a powerful marker of change of topic. For example, S8 falls at one of the major topic changes in the whole text, where no fewer than five signals, both extrinsic and intrinsic, coincide, but in spite of this the author chooses not to start a new paragraph at this point.

Passage B only has two short sentences, and one of these begins a paragraph. A glance at Figure 20 will show that this writer does not vary his sentence length to the same extent as the others, tending to write a succession of average length sentences of approximately one T-unit each. The short sentence beginning the paragraph at S35 is worth examining. The paragraph here, while acceptable, is nevertheless surprising. The writer has placed a short bridging paragraph of two sentences only (S33 and S34) between his description of two different methods of definition, but this paragraph could quite logically have been added to either of its neighbours. It is as if the writer wanted to make doubly sure that the reader was aware of what he was doing, and may have been a way of emphasising the importance of the final improved definition considered in the last paragraph.

2.1.3 Contrast words If short sentences do indeed indicate transitions or contrasts, then it is not surprising to find short sentences and contrast words in combination. Of six examples of contrast words present here, (two in each passage), no fewer than five occur as part of a short sentence or immediately following one.

Traditional teaching on rhetoric and style has always insisted that no sentence should begin with 'and' or 'but', but this 'rule' flies in the face of actual practice. Two sentences in Passage A - 'But Nature has also provided certain defensive forces' (S9) and 'But he has also exposed the soil to the action of the atmosphere' (S24) as well as one in Passage B - 'But there are difficulties here too' (S13) are examples. In all three cases the contrast is not one between adjacent sentences, but between adjacent information blocks : the 'but' represents a linkage at a hierarchically higher level, at the level of the macrostructure (Van Dijk, 1977), and would be referred to by Van Dijk as a 'macro-connective'. The traditionally proscribed 'but', in contrast, links two independent clauses at sentence level.

Other contrast words and phrases which could have been used include 'nevertheless', 'in spite of this', 'despite this', 'on the other hand', 'in comparison', 'in contrast', 'conversely', 'notwithstanding', etc. These include markers of concession as well as of contrast. It seems likely that 'but' and 'however' are the most common.

2.1.4 Text reference signals (ie reference signals which refer to chunks of text rather than to single propositions). These may take the form of phrases, clauses or complete sentences. One of their functions is to indicate to the reader that a new chunk is on the way (cataphoric signals) or that one has just passed (anaphoric signals). Such signals would be considered by Van Dijk as referring to macro-propositions within the macro-structure of the text. It should be noted that the terms 'cataphoric reference' and 'anaphoric reference' are being used here rather more broadly than is usually the case.

2.1.4.1 Cataphoric signalling is a more useful indicator of text chunking for the reader, in that it prepares him for what is to come rather than simply informing him that an information block has just passed. Reference to the first person singular (eg I shall now turn to . . .) or to the first person plural (eg Let us now consider . . .) is a common metalingual device serving to focus attention on a particular point, very often a new topic. A cataphoric text signal is frequently part of a topic sentence.

The largest number of examples is to be found in Passage B - in fact, cataphoric text reference is the typical method for this writer to introduce a change of topic. He only uses first person reference in the first sentence (Let us consider . . .), but liberally uses non-

personal cataphoric reference throughout the text, culminating in two enumerations, introduced by 'First' (S25, S36), 'Second' (S27, S37), 'Third' (S39) and 'Finally' (S30).

The writer of Passage A uses text reference more sparingly, but introduces first person reference twice, on both occasions at points where he also begins a new paragraph. These are 'We may sum up by saying that ...' (S18), and 'Now let us consider what happens...' (S20). Apart from these we could have included 'For example...' (S29), as this phrase introduces an extended example which continues right up to the end of the passage. It was decided, however, that as connectors such as this have a subordinating function, and usually signal movement at a lower level than those previously mentioned, for the purpose of this analysis they would be omitted. The problem here is that certain text structuring devices can be used to relate items at all levels, from the proposition to large chunks of text.

The writer of Passage C, in contrast to the other two, uses no cataphoric references of any kind.

2.1.4.2 Anaphoric signalling Anaphoric text signals, together with paragraphing and topic sentences, help to indicate the hierarchical organisation of the text. For example, in Passage A 'thus' (S16) helps to combine the two preceding information blocks, and in the same passage the anaphoric reference of the two-way signal 'We may sum

up by saying...' (S18) combines the preceding four chunks into one superordinate block.

Passage A makes the greatest use of anaphoric signals. The main weakness of their use, however, is that while they specify that a block has been completed they do not indicate where the block began. For example, the very strongly marked transitional sentence 'All this is natural' (S8 in Passage A) closes an enumeration of erosive forces and their effects, but the reader may not have realized immediately at the beginning of the block that he was being presented with a list. Anaphoric signals, therefore, are not as helpful for the reader in marking text chunking as are cataphoric signals, and they have been omitted from this analysis.

2.1.4.3 Two-way signalling This occurs when one phrase or sentence points both backwards and forwards, and is thus the most powerful type of text reference signalling. In Passage B S22 begins with the phrase 'As a result of this problem . . .' which refers back to the problem which has been considered throughout the whole of the first part of the passage and forward to the second half, which will discuss the result. It is thus the pivot upon which the entire text turns, and a misunderstanding of the reference would seriously impede an understanding of the whole text.

In Passage A the phrase 'We may sum up by saying . . .' (S18) serves a similar function, though at a lower

hierarchical level, by summing up the whole of the first part of the passage before a major transition to the second half. This particular signal points back to the part of the passage that is to be summed up, and forward to the summary itself. Like all other anaphoric signals, however, it is difficult for the reader to be completely certain as to where the chunk being signalled begins. The author's paragraphing would suggest that it should be the whole of the passage up to S18, but if we consider meaning it is more likely to be the section from S8 or S9 only.

2.2 Intrinsic signalling Apart from verb tense change, intrinsic signalling presents some problems of identification if anything more than intuition is to be used. Reference should be made to Table 22 for a complete breakdown of all signals.

2.2.1 Verb tense change Change of verb tense is suggested by Koen, Becker and Young (1968 : 174) as one of the patterns of formal markers of information chunking within the grammatical system. It is doubtful, however, how far analysis according to verb tense can be successfully applied to expository prose, although it is probably suitable for narrative. Change of time may be more important than change of aspect as can be demonstrated by some rather odd transitions marked between the simple present and present perfect tenses in Passage C.

The analysis works well in Passage A, but this is the only passage which includes a narrative, nicely

demarcated by a switch to simple past tense (S29). A switch from general truth to completed action also supplies an effective boundary (S23). In Passage B, a switch from simple present to simple past and then back to simple present (S24, S25), and again from present perfect to simple past (S35) are similarly valid. In Passage C, change of aspect occurs rather than change of tune, producing an intuitively less happy result.

2.2.2 Lexical cohesion chains (See Appendix 5a) We are here concerned with the repetition of reference from one sentence to another through particular sections of the text. The lexical cohesion chain is similar to Halliday and Hasan's 'cohesive chain' (1976 : 15) and Young and Becker's 'Lexical equivalence chain' (1966), except that it is not necessary that there should be only one presupposed element : the whole range of possible lexical relationships may be involved, including collocation. Thus, in Passage B the following items contribute towards forming a chain : urban settlements, the continuous urban sprawl that is Greater New York, Melbourne, cities, the legal city, parts of the urban area, places with a few hundred people, settlements, and contiguous cities. Parallel chains of the type described by Halliday and Hasan also occur in Passage A, where pronominal reference to 'man the pioneer' can be followed through five consecutive sentences.

In analysing these texts into lexical cohesion chains, three main problems arose. The first concerned

the weighting to be given to different parts of the sentence. It seemed unreasonable, and in any case would have been too unwieldy to handle, if every part of the sentence had been made equally important for the choice of reference items. It was therefore decided to discount subordinate clauses entirely, and to consider only noun phrases within or very closely linked to the main clause. This would include complements and restrictive adjectival clauses, but exclude non-restrictive adjectival clauses and adjuncts of time and place (except for the occasional obligatory adjunct of place expressing range).

Items selected were sub-divided into main references and subordinate references. Only main references could begin a chain, a main reference being the headword of a noun phrase as selected above. A subordinate reference could be any other part of the noun phrase, and also occasionally a main verb, though a verb could never begin a chain.

The second and third problems were linked, in that both were concerned with distance - distance of references from each other within the lexical system and physical distance within the text itself.

The matter of distance within the lexical system could only be solved by intuition and common-sense. Obviously there could be no question about the exact repetition of a lexical item, but on other occasions interpretation was necessary. For example, the chain in Passage A headed 'Erosive forces of nature' was taken to

include the rain, the frost, force of gravity, alternate heating and cooling, and the wind, although at no point does the author explicitly state that these are erosive forces. This becomes more problematic in Passage B, where it has to be decided whether to include in one chain, that of 'Physical structure', a continuous distribution of housing, population above a certain density, and the intensity of traffic. The author clearly sees these as being closely related, the latter three being hyponyms of the former, but it seems questionable whether the 'physical structure' of a city can include continuously changing features only revealed by statistical analysis. Obviously there can be no clearcut answer. Cohesion depends on relatedness, and there can never be an absolute solution. If there is no exact repetition of a lexical item, or substitution by a pro-form, then it is a matter of 'more or less', and judgment on this is inevitably selective.

Distance within the text itself is a simpler matter, however. The cohesive effect becomes weaker as physical space increases, and for the purpose of a particular analysis an exact cut-off point can be chosen. In this case, items were traced through T-units rather than sentences. Three consecutive references were considered necessary at the beginning of every chain, after which a gap of up to two T-units was allowed, the chain being closed only if no more references followed up to the end of the passage. If a gap of six T-units existed, followed

by a further three or more consecutive references, then a new boundary could be marked, even though the old chain could not be closed.

This is a fairly conservative method of analysis, but it nevertheless results in between seven and eleven boundaries in each passage. Of the three passages, only C does not show a more or less complete succession of cohesion chains from the beginning of the text to the end. Where gaps appear, however, this does not necessarily mean the absence of lexical cohesion, but instead that according to the strict rules applied in this analysis the boundaries of cohesion chains cannot be marked. Thus in Passage C the large gap between TUs 20 and 34 contains at least one discernable semi-chain, notably H, 'The climax type of forest' which is almost certainly identifiable by the reader.

2.2.3 Movement from particular to general (See Appendix 5b)
The work of Christensen and Koen, Becker and Young suggest that a new paragraph often begins where there is a movement from a more specific to a more general level, but whereas this is usually intuitively clear, it is difficult to pinpoint in an objective manner.

In this analysis, nouns and noun phrases have been examined according to the same criteria as for the lexical cohesion chains described above, and where occurring in the same semantic field have been hierarchically organised according to their hyponym/superordinate and whole/part relationships (See Lyons, 1977 : 291-301).

The relationships have been set up and verified in the following manner. If x is the hyponym and y the superordinate, or x the part and y the whole, then for each occurrence the following questions can be asked : either, 'Is x a kind/sort/type of y?' or, 'Is x a part of y?' If the answer is positive, then the item represented by x can be inserted beneath the item represented by y in a tree diagram. Verbs and adjectives may also be treated in a similar manner, though only noun phrases were considered here. In the case of all three passages, the resultant representation is a more or less developed tree, but it seems likely that a series of independent strings might be a more common realization.

In the indication of upward movements, care should be taken not to jump across the tree in such a way that the two juxtaposed items are too distantly related semantically to represent a particular/general contrast. For example, in Passage A a movement from sphagnum at the bottom of the 'vegetation' node to the force of gravity, dominated by 'the erosive forces of nature' node would not be viable : this would be more likely to indicate a movement to a new lexical cohesion chain than a movement from particular to general. The rule should therefore be followed that only movement up the tree through successive nodes, and not movement across the tree to parallel branches is allowed. This rule has been broken in one place in Passage A, which is represented by the most complicated and diversified tree. Here a movement has been allowed from TU9/10 to

TU11, ie from the wind (an erosive force) to certain defensive forces. It is likely that cross-movement would in fact be possible at the very top of the tree, and here such a close relationship exists between the two nodes - both represent forces of nature - than an exception has been made.

Again, Passage C has the smallest number of boundaries, with only three, compared to four for Passage B and five for Passage A. This is the inevitable result of Passage C's less well-developed tree, and continues the by now well-established pattern.

3. Grimes' Semantic Grammar of Propositions as a model for text chunking A number of psychologists and psycholinguists have suggested methods for the diagrammatic representation of the information structure of a text, eg Fredericksen (1972) and Crothers (1972). The resulting models, however, are inevitably influenced by the differing purposes of the researchers, and not all can serve as tools for the discourse analyst. For example, Crothers is interested primarily in the structure of the passage as it is represented in the reader's memory. He therefore treats the original organisation of content by the author as superficial, imposing his own organisation on the concepts he identifies in the text.

A model more closely linked to the author's own organisation of the text is that worked out by Grimes, and based on his Semantic Grammar of Propositions (1975).

Grimes is more interested in the structure of the text itself than in reader memory, but his model was taken up by Meyer (1975), slightly modified, and used for a series of experiments on the effect of the organisation of text on recall. Meyer's work carries implications that Grimes' model might throw some light on paragraphing, for she found that a short cut towards reconstructing the top levels of the content structure of the text was to look at the chunking of information into paragraphs, and particularly at the so-called 'topic sentences'. Grimes himself also mentioned in passing (Grimes 1972 : 240) that paragraphing could be a result of a change of depth in the tree structure of the passage. It was therefore considered that there might be some relation between paragraphing (and therefore text chunking) and movement from a lower level to a higher level in the tree.

In order to investigate whether this model could indeed throw any fresh insight onto this problem, it has been applied, with some slight modifications, to Passages A, B and C. The resulting hierarchical tree diagrams are represented in Appendix 6.

3.1 Grimes Semantic Grammar of Propositions : description

The application of the Semantic Grammar to a text results in a hierarchical tree, revealing the content structure of the passage, and consisting of layers of propositions. (A proposition, according to Grimes, consists of a predicate plus its arguments). Two types of relation are present

in the model, role relations and rhetorical relations, the latter being the most important for the overall organisation of the text, and tending to occur most often at the higher levels.

Role relations reveal the way in which lexical predicates, usually verbs, relate to their arguments, which are based on Fillmore's cases (Fillmore, 1968). The nine role relationships recognised by Grimes are those of agent, instrument, force, vehicle, patient, benefactive, latter, former and range. These will not be considered in detail here, as they are of little significance for higher level organisation.

The more important rhetorical relations reveal the ways in which ideas relate to each other. A rhetorical predicate may have whole propositions, either lexical or rhetorical, as its arguments, so tends to be found at the top levels of the tree. There are three types of rhetorical predicate, their classification depending on the relationships of coordination and subordination. These relationships, however, are rhetorical and not grammatical, so that a realization of a rhetorical relationship of, for example, hypotaxis, may not be a grammatically subordinate clause. Thus some of the insights of Christensen's unformalised theory of paragraph formation and structure may perhaps be made more explicit within the framework of Grimes' Semantic Grammar.

The three main types of rhetorical predicate as identified by Grimes are paratactic, hypotactic and neutral,

the neutral predicates being either paratactic or hypotactic depending on the emphasis placed on them by the author. Grimes' eighteen rhetorical predicates can be summarised as follows :

A. Paratactic

(i) Alternative This gives a list of mutually exclusive options, which may be expressed in the text as 'either . . . or'.

(ii) Response This very often occurs at the highest level, and consists of Question/Answer, Problem/Solution or Remark/Reply. (It seems that more possibilities could be added, eg Contrast (Positive)/Contrast (Negative).)

B. Hypotactic

(i) Attribution relates a description to what it is describing.

(ii) Equivalent restates an idea in a different way.

(iii) Specific relates more specific information to a general statement.

(iv) Explanation explains previously stated information.

(v) Evidence to support an idea. This probably includes exemplification.

(vi) Analogy to support an idea.

(vii) Manner shows the way an event is performed. It tends to be a lower level predicate than the rest.

(viii) Adversative relates what did not happen to what did happen.

(ix) Setting Time gives the time of the setting at which the information being related occurred. This and the next two predicates are commonest in narratives.

(x) Setting Location gives the location of the setting at which the information being related occurred.

(xi) Setting Trajectory gives the changing background of location and time when the character in a narrative moves from place to place.

(xii) Representative Identification singles out one element in a group and makes it stand for the group as a whole.

(xiii) Replacement Identification makes one thing stand for something else.

(xiv) Constituency Identification identifies a part in relation to a whole.

C. Neutral

(i) Collection presents a list of related elements, which may be concurrent or consecutive. If consecutive, they represent a time sequence of events, which Meyer suggests should be related by means of indices for time using Litteral's time topology (Litteral, 1972).

(ii) Covariance expresses condition, result or purpose, with one argument as the antecedent and the other as the consequent.

3.2 Grimes Semantic Grammar of Propositions : application

Although the analysis results in a tree structure, even Grimes himself does not claim that all text is hierarchical. In fact, he suggests that paragraphing may be the result not only of a change of depth in tree structure, but also of 'a shift in theme in a uni-dimensional text' (Grimes, 1972 : 240). However, certain types of writing, eg explanation and argument, are undoubtedly hierarchically organised, and for such texts this kind of analysis seems suitable.

Meyer suggests that analysis can begin either at the bottom of the tree or at the top. Analysis from the bottom upwards is easier but more time-consuming, and results in a more detailed model than may be necessary. As for the purpose of the present analysis only rhetorical predicates were considered, analysis proceeded simultaneously from the top and the middle (ie at the level of the complete clause rather than the predicate plus its arguments).

While the identification of the lexical predicates with their related arguments is relatively objective, problems arise with the rhetorical predicates. Rees and Urquhart (1976) point out that all discourse analysis is of necessity largely intuitive, as many of the inter-sentential relationships ofthe normal text are implicit. Thus there are almost as many possible interpretations as there are readers, and ideally the writer himself should make the analysis. In these particular texts there are few explicit inter-sentential connections, although some of the larger blocks of information are overtly signalled.

The higher up the tree the analyst moves, the more subjective his judgment inevitably becomes. The problem is particularly great with this type of text, being extracts from longer passages, so that much of the evidence needed to link them together at the highest levels is missing, and has to be inferred. This is especially noticeable with Passage C. Indeed, it is almost impossible to find authentic <u>complete</u> text of a manageable length for analysis, except of course for newspaper articles, or such abnormal texts as abstracts, or students' essays. The dangers of excessive subjectivity can be lessened, however, if analysis proceeds from the bottom upwards rather than from the top downwards.

More specific, though minor, difficulties were also present. Although Meyer claims to have successfully applied Grimes' model to similar expository texts, as far as these passages were concerned a few modifications in the predicates themselves were found necessary. These were as follows :

(i) All Collection predicates were considered as being paratactic, and all Covariance predicates as being hypotactic.

(ii) Covariance predicates were increased in number, so that they included Cause as well as Result, Purpose and Condition.

(iii) Two extra hypotactic predicates were introduced, Time and Amplification. (Jones and Faulkner,1961). The Amplification predicate was necessary when extra information was added to a statement which was neither more specific, equivalent in meaning to that statement, nor an explanation of it.

Basically, however, the labels used are not of great importance in themselves in this analysis, and from now on they will be ignored.

3.3 The Vertical Movement Index All movements from a lower to a higher level in the hierarchy have been marked in Appendix 6 by arrows. There are 13 such movements in Passage A, 12 in Passage B and 9 in Passage C. They range from a massive leap of five levels to Level 1 in A, and 7 to Level 2 in B, to much smaller and less significant jumps of only one or two levels lower down in the tree.

In order to correlate the vertical distance with the level reached, and to give extra weighting to the higher levels, the Vertical Movement Index (VMI) was devised. This index can be represented by the formula

$$\frac{\text{Movement} \times \text{Level}}{100}$$

Movement is the number of levels through which movement takes place.

Level is the level finally reached, and for this it is assumed that there are ten possible levels, weighted from 10 at Level 1, through nine at Level 2, eight at Level 3, etc. to 10 at Level 1. (The levels are numbered from the top of the tree downwards). Thus, a movement from Level 5 to Level 2 is represented by the Index as $\frac{3 \times 9}{100} = 0.27$. The maximum possible score would be 1.00 for a movement from Level 11 or below to Level 1. The maximum possible score on these particular passages would be 0.60 on Passage A, 0.90 on Passage B and 0.70 on Passage C. The various weightings

given to the upward movements in the three passages are shown in Table 23. A movement is only measured when it actually occurs in the text between one sentence and the next : thus some of the attributive predicates which occur in the middle of other higher level predicates are not considered.

TABLE 23

VERTICAL MOVEMENT INDEX MEASUREMENTS FOR EACH PASSAGE

PASSAGE A			PASSAGE B			PASSAGE C		
SENTENCE	20	0.50	SENTENCE	22	0.63	SENTENCE	11	0.40
	27	0.27		33	0.35		8	0.24
	12	0.16		11	0.27		13	0.18
	18	0.16		17	0.16		24	0.16
	16	0.14		8	0.14		21	0.14
	9	0.09		39	0.12		16	0.09
	23	0.08		9	0.07		22	0.08
	24	0.08		15	0.07		27	0.07
	17	0.07		18	0.07		30	0.04
	26	0.07		27	0.07			
	32	0.07		5	0.06			
	36	0.06		29	0.05			
	33	0.05						

4. Experiment 3

Most educated native speakers have an almost intuitive knowledge of where it is or is not appropriate to begin a new paragraph. This may have been directly and consciously taught to them as part of the educational process they passed through, but is at least partly the result of a lifetime of reading English prose. Previous experiments (eg Koen, Becker & Young (1968)) indicate that considerable agreement is reached between such subjects when

asked to insert paragraph boundaries into a text, and that where differences arise these can be resolved to most subjects' satisfaction by mutual discussion. It was decided, therefore, to use such subjects as 'native informers', and to try to relate their decisions over the structure of the texts to the signals already isolated by the two types of text analysis described above. The following hypotheses would be tested :

Hypothesis 3

A paragraph or information block boundary is most likely to occur under either or both of the following conditions :

(a) Where there is a movement upwards to a higher node in the hierarchical structure of the text. The greater the vertical distance of the movement, and the higher the node, the more likely is it that a new paragraph or information block will begin.

(b) Where there is a movement downwards or across the tree to a node which dominates a number of propositions through at least two branches. The more propositions the node dominates, the more likely it is that a new paragraph or information block will begin.

Hypothesis 4

Extrinsic signals are more effective markers of paragraph and information block boundaries than are intrinsic signals.

4.1 Subjects and method For this experiment, seventeen volunteers from Moray House College of Education, Edinburgh, were used as subjects, drawn from the Departments of Geography, Biology, Chemistry, Sociology, Economics and Modern Studies. Three further volunteers from the Departments of Geography and Biology at the University of Malawi were added, to make a total of twenty subjects.

Each subject was presented with the three passages, A, B and C, with the authors' paragraph boundaries removed, and were asked to mark the places in the texts where they would judge paragraph boundaries to be appropriate. A difference between this and similar experiments, notably those of Koen, Becker and Young (1968), is that here subjects are required to paragraph hierarchically, first dividing the passage into two paragraphs only, then dividing each paragraph into further sub-divisions, until finally segments below the level normal for an orthographic paragraph are reached. In all, three succeeding divisions were made of Passage B and four of A and C. The detailed instructions given to the subjects are reproduced in Appendix 7, and these should now be referred to.

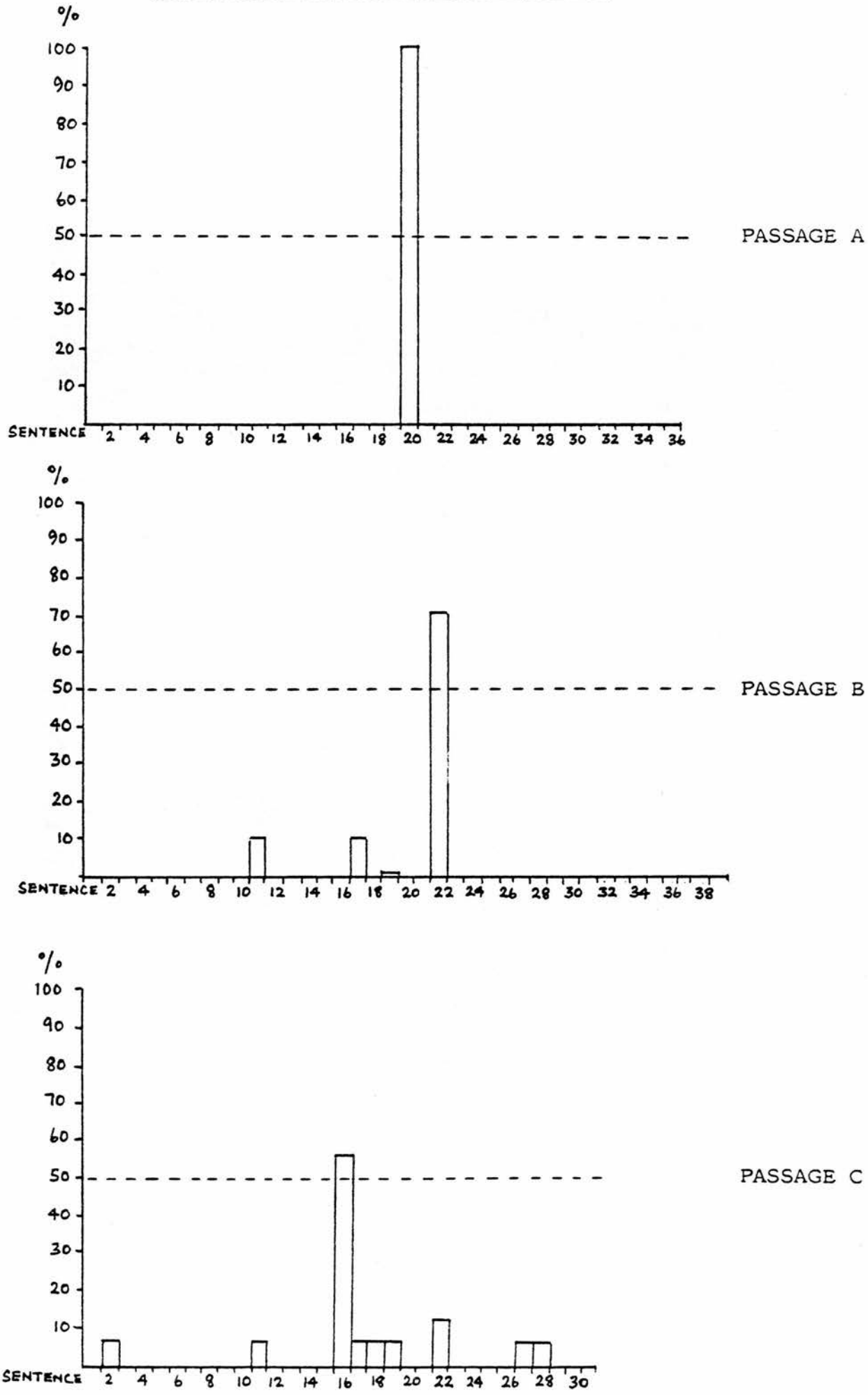
As in Experiments 1 and 2, the subjects used were experienced mother-tongue readers of the type of English exemplified by the texts - each was a graduate teacher at tertiary level in a subject area in some way related to the area of geography found in at least one of the texts. Teachers of English and Linguistics were again avoided as subjects, under the assumption that they would have too many preconceptions about text structure.

Of the twenty subjects, nineteen made completed returns, and of the fifty-seven paragraphed texts received two were spoiled. Thus the final total of returns was fifty-five : nineteen for Passage A, and eighteen each for Passages B and C. An analysis of these returns can be seen in Table 22, where boundaries chosen by 50% or more of the subjects for Divisions 1 and 2 and 40% or more for Divisions 3 and 4 are placed against the boundaries suggested by the analyses of observable signals and the VM Index. These tables should be used in conjunction with the report of the statistical analyses given in section 4.3.

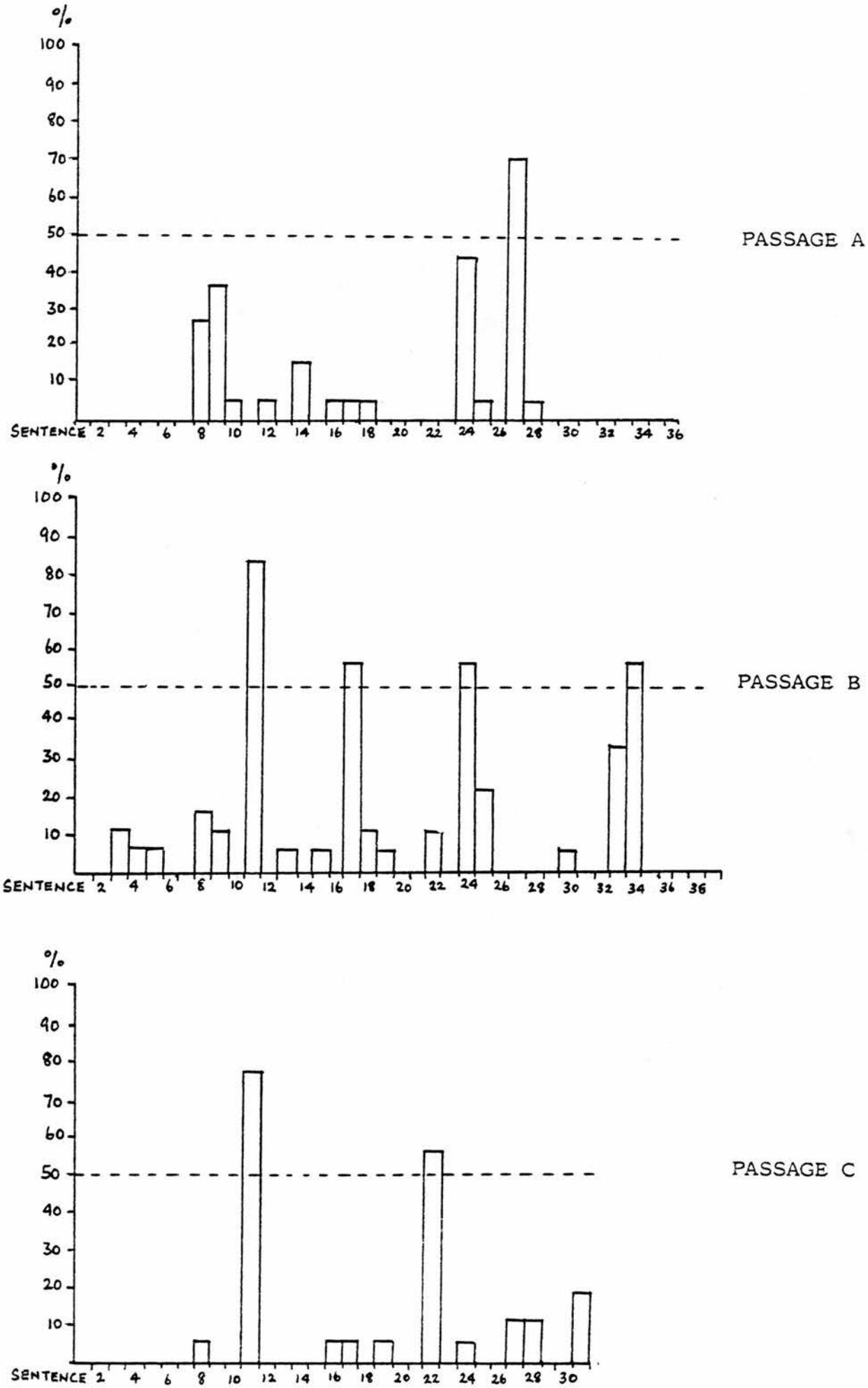
4.2 Raw results The results of Experiment 3 are represented as histograms in Figures 21 and 22. The figures themselves can be found in Appendix 7. From Figure 21 it can be seen that in Passage A every subject without exception chose S20 for the first major division. In Passage B, 72.22% chose S22, and in Passage C 55.55% chose S16.

When we look at the second divisions, however, the choices for Passage A are less uniform, with 68.42% for S27, and 36.84% for S9. If, however, S8 and S9 are combined, with S8 considered as a bridge sentence, the percentage rises to 63.16. Passage B has only one very high score, 83.33% for S11, but there are suggestions of at least one bridging element : if, for example, the number of subjects choosing S33 and S34 are combined, the total

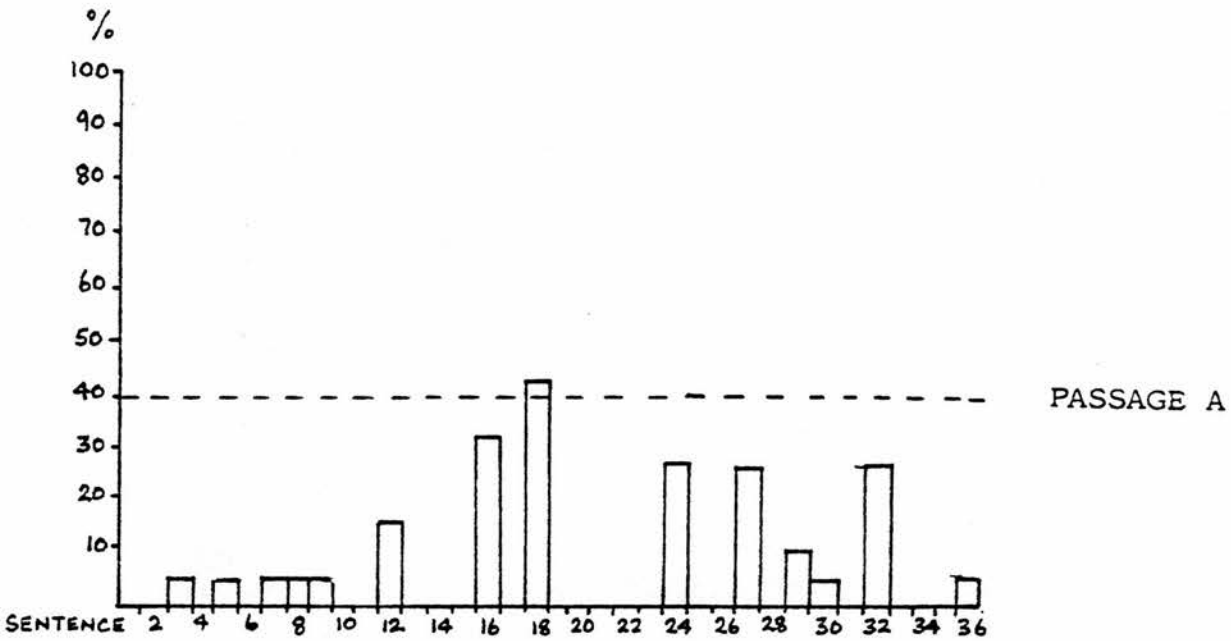
FIGURE 21 EXPERIMENT 3 : RESULTS

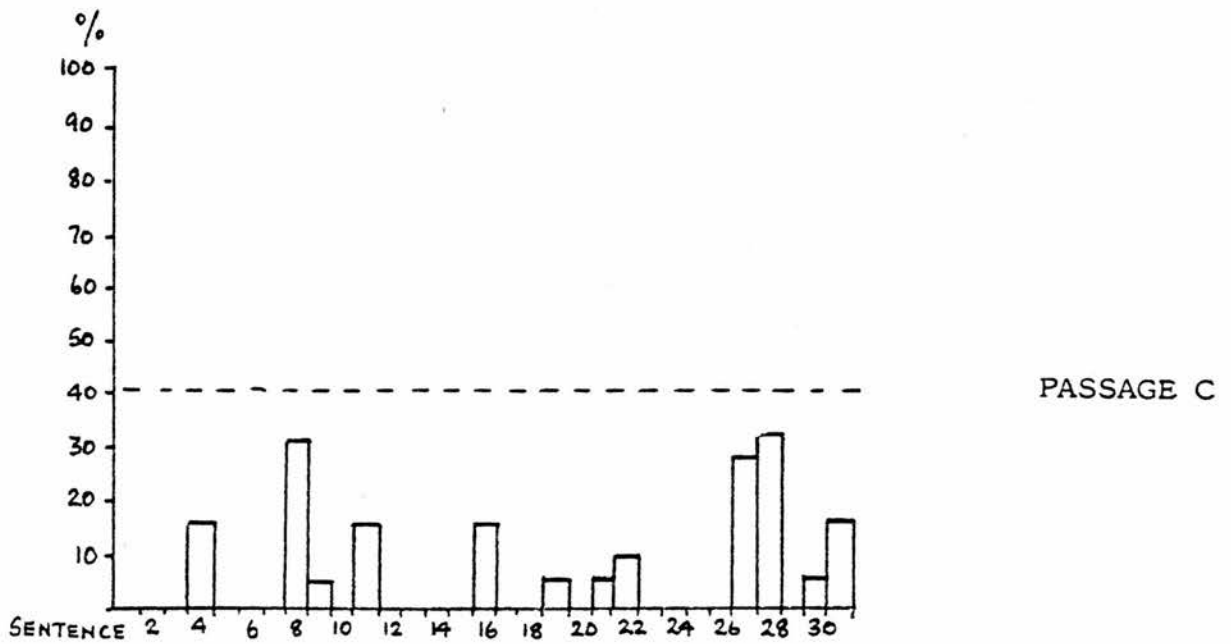


(a) BOUNDARIES CHOSEN AT DIVISION 1

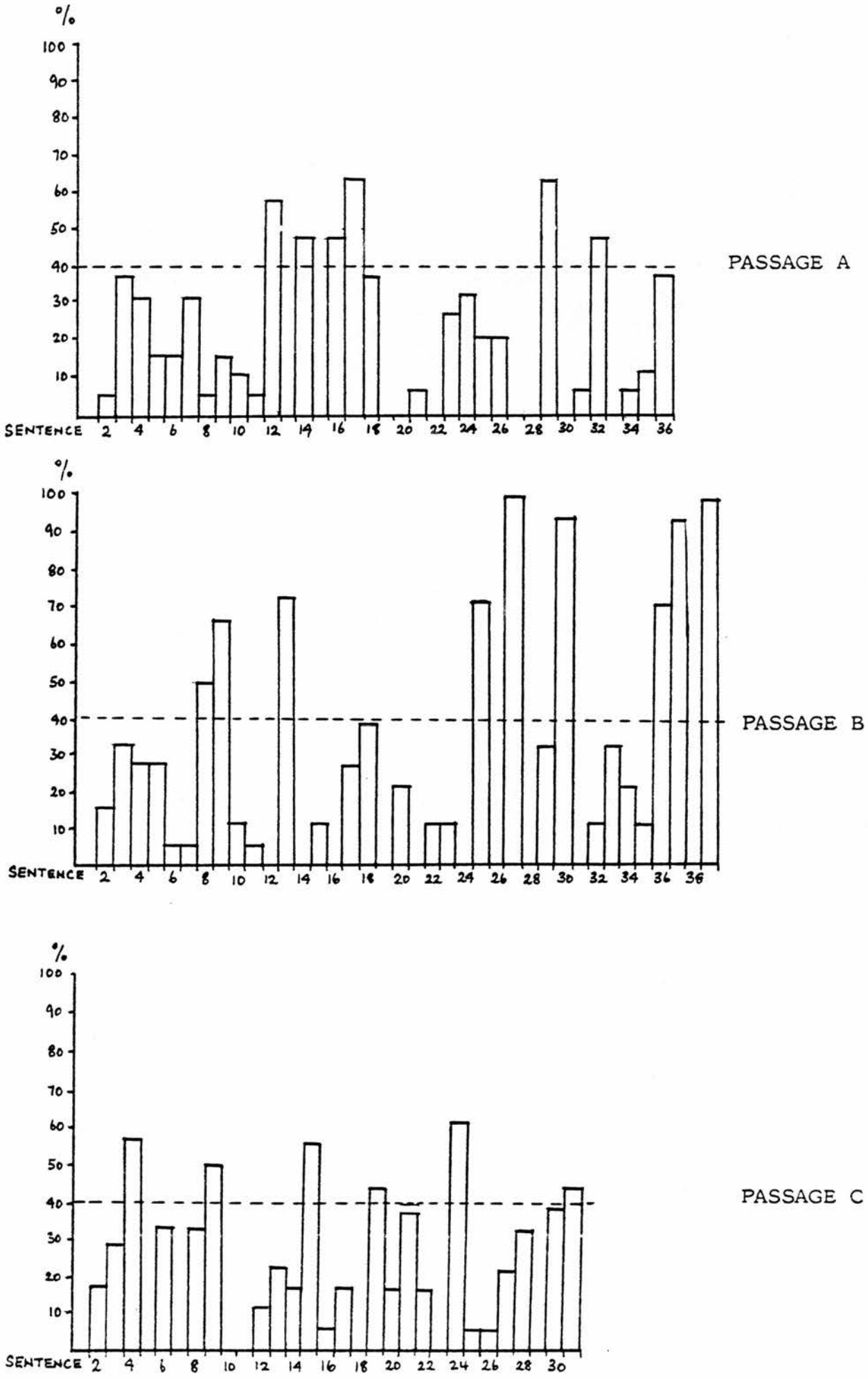


(b) BOUNDARIES CHOSEN AT DIVISION 2





(c) BOUNDARIES CHOSEN AT DIVISION 3



(d) BOUNDARIES CHOSEN AT DIVISION 4

percentage rises to 88.89. Bridgings and transitions are considered more fully in Chapter 10. Passage C has the most clear-cut choices at the second division, with 77.78% for S11 and 55.56% for S22.

The third 'paragraph' division was only required for Passages A and C, and this shows much less agreement among the subjects. No sentence is chosen by more than 42.11% of the subjects in Passage A (S18), and in Passage C no sentence is chosen by more than 33.33% of the subjects.

The closest agreements on the boundaries of the smallest segments at the fourth division is found in Passage B, where eight boundaries are chosen by over 60% (Ss 9, 13, 25, 27, 30, 36, 37, 39). S27 and S39 were chosen by every subject. In Passage A, however, only Ss 17 and 29 exceed 60%, and in Passage C S24 alone reaches this figure.

Although Koen, Becker and Young treated paragraph boundaries chosen by 20% or more of their subjects as significant, it was considered here that this figure would be too low, especially seeing that fewer subjects had been used than in Koen, Becker and Young's experiments. It was therefore decided to analyse only those boundaries chosen by 50% or more of the subjects at Divisions 1 and 2, and by 40% or more at Divisions 3 and 4, a total of 33 boundaries.

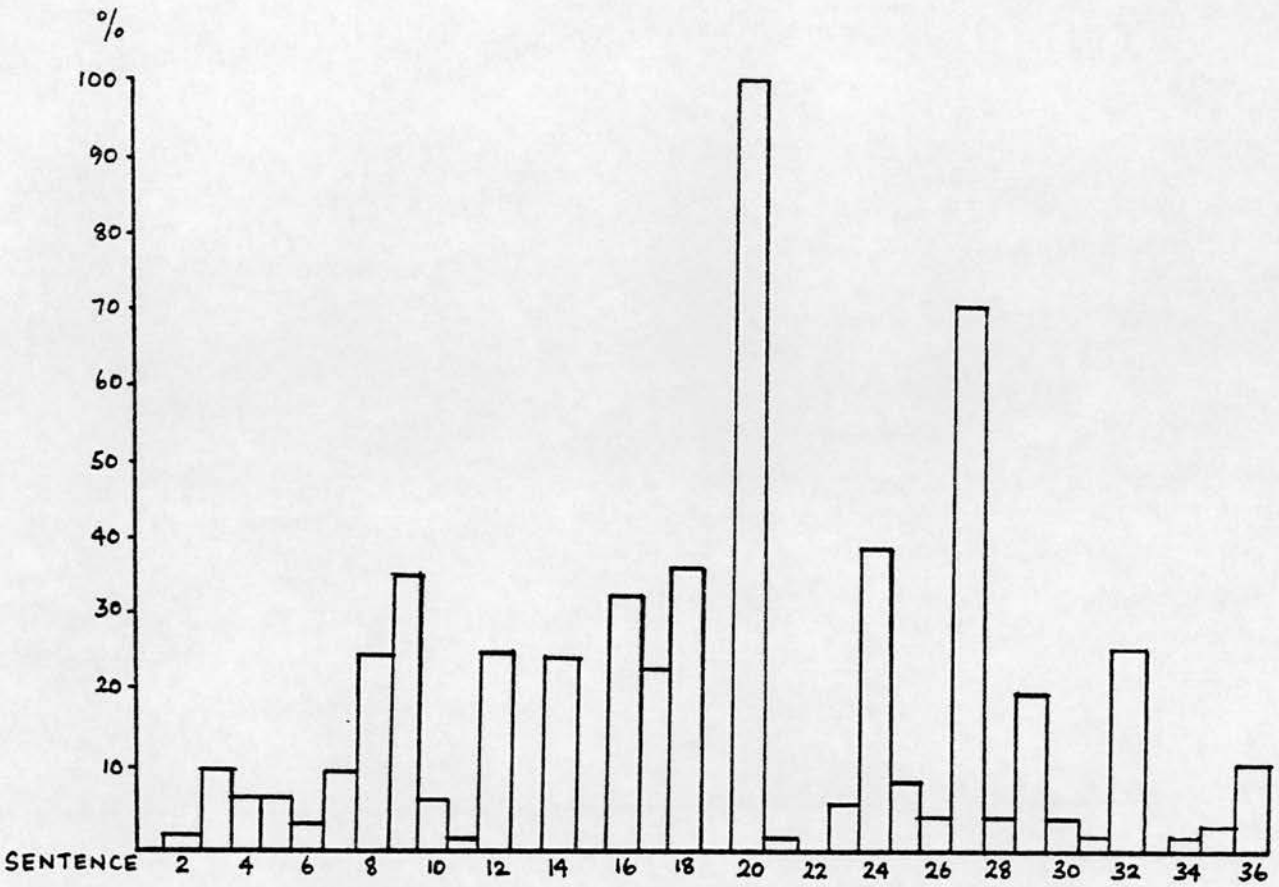
This decision has, however, one major drawback : as the percentage of choices is in every case taken at one level only, a few boundaries chosen by a large number of

subjects but at different levels fall below the required figure when each level is considered separately. For example, S24 in Passage A receives 21.05% at Division 2, 26.32% at Division 3, and 31.58% at Division 4. Other sentences with a similar pattern are S33 in Passage B and S28 in Passage C. It should be stressed that such sentences are exceptional : in most cases there is no doubt as to which division a sentence belongs to.

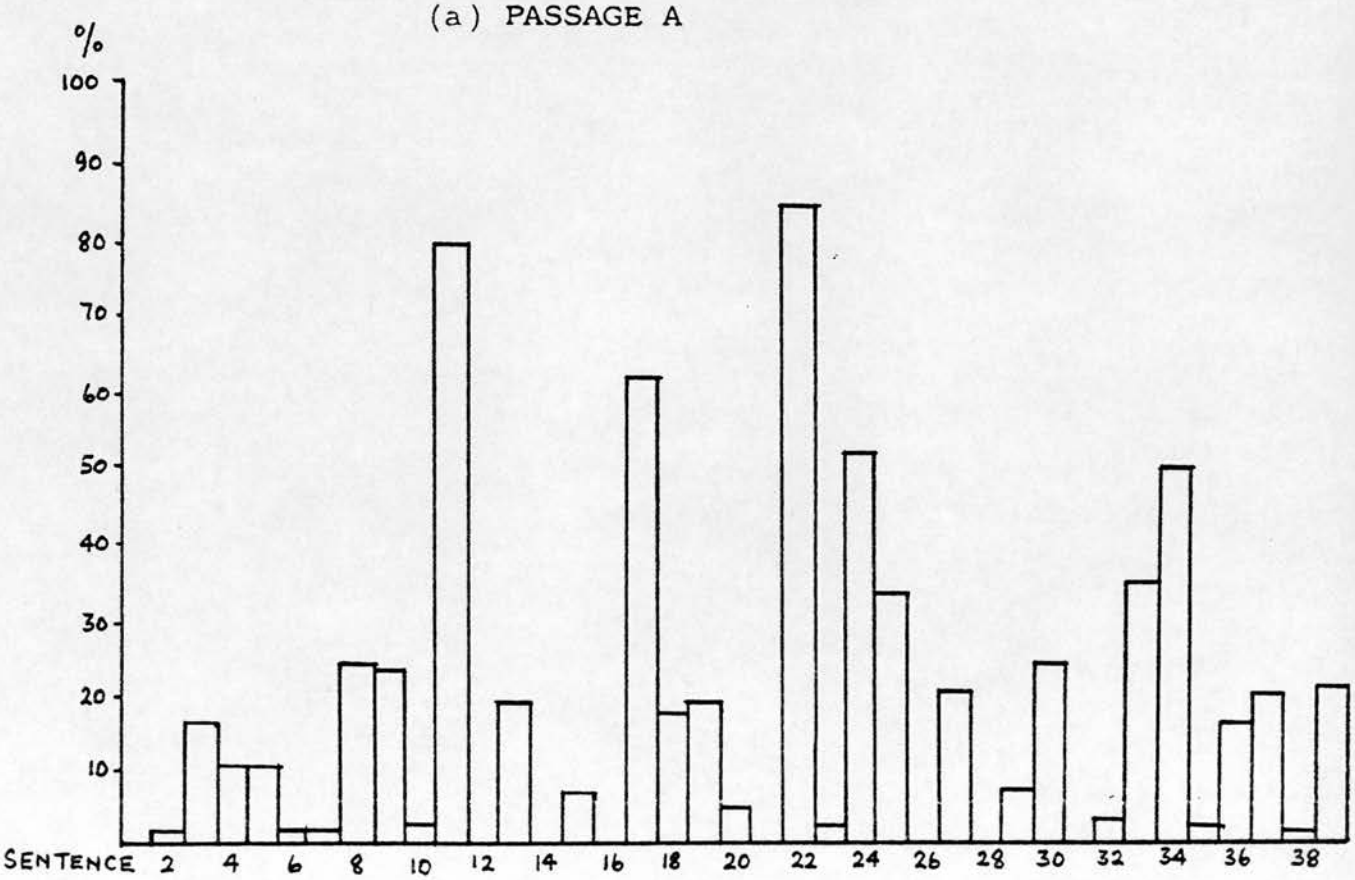
Figure 22 illustrates how this problem was tackled. Here, the choices at all four divisions are combined by weighting each choice, the higher the division, the greater the weighting allocated to it. Thus, a total for each boundary chosen was calculated, by allotting five points for each choice made at Division 1, four at Division 2, three at Division 3, and one at Division 4. Taking the maximum possible total for each possible boundary as five times the number of subjects (ie 95 for Passage A, 90 for Passage B and 90 for Passage C), the total score was in each case converted into a percentage. Any boundary then obtaining more than 40% of the possible total, and which had not been selected according to the first method, could be added to the 33 already chosen for analysis. Perhaps surprisingly, the only boundary fulfilling these criteria was S28 from Passage C, so this was added to Division 3, making a final total of 34 boundaries.

FIGURE 22 EXPERIMENT 3 WEIGHTED CHOICES

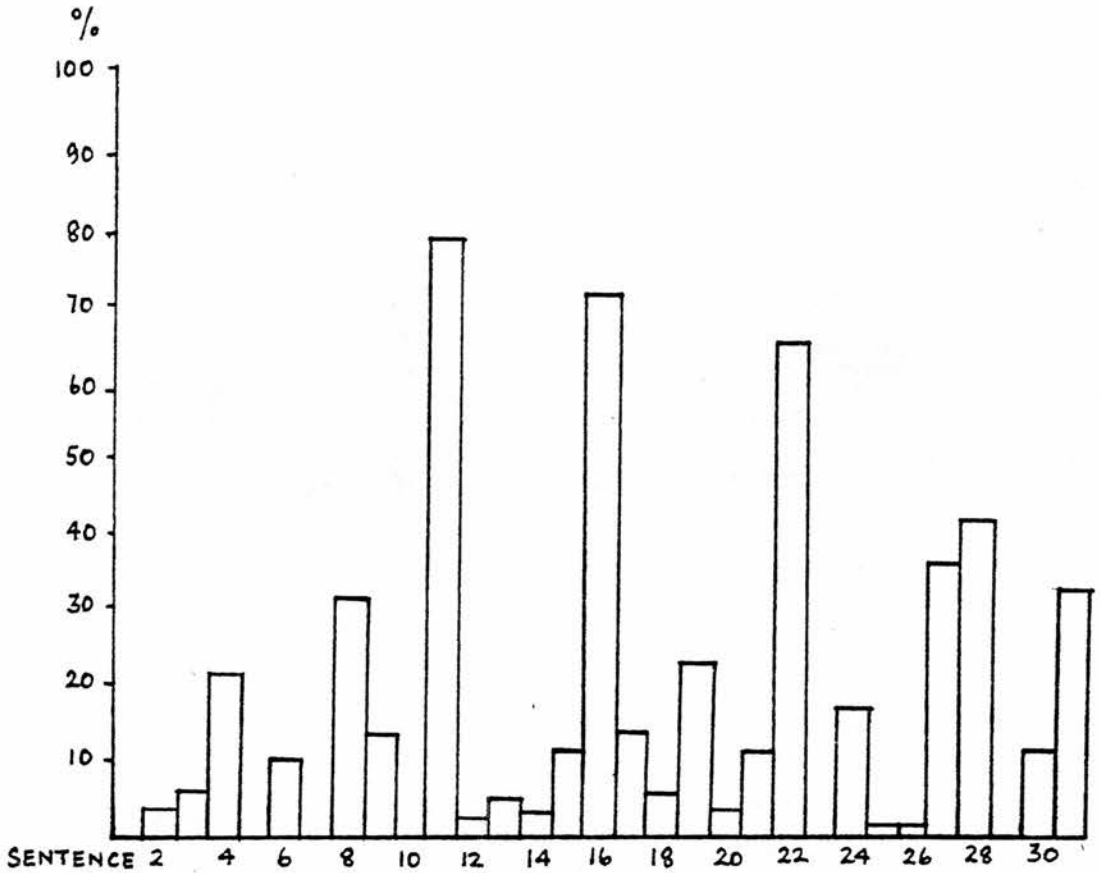
(ALL DIVISIONS COMBINED)



(a) PASSAGE A



(b) PASSAGE B



(c) PASSAGE C

WEIGHTING

DIVISION	1	=	5
	2	=	4
	3	=	3
	4	=	1

4.3 Statistical Analysis Statistical analyses relevant to the two hypotheses will now be made. In all cases a probability level of 0.05 is required to indicate significance, while a level of 0.01 would indicate high significance.

4.3.1 Spearman Rank Correlation Coefficient : r_s (1) (See Appendix 8). For this test, only the eighteen boundaries selected by over 40% of the subjects which were also marked by an upward movement, however small, were chosen.

Null Hypothesis (H_o) : There is no relationship between the VMI measurements allotted to the boundaries and the divisions at which they were selected.

Alternative Hypothesis (H_1) : There is a significant positive relationship between the VMI measurements allotted to the boundaries and the divisions at which they were selected.

After applying the correction factor for ties (T) to both variables, an r_s of 0.61 was obtained, thus indicating that a relationship does, in fact, exist. This relationship is highly significant, ($p = < 0.01$), so the Null Hypothesis can be rejected.

4.3.2 Spearman Rank Correlation Coefficient : r_s (2) (See Appendix 8). For this test, the 34 sentences beginning after an upward movement, however small, were chosen. The second variable was the weighted percentage combining subject choice for all divisions (Figure 22).

Null Hypothesis (H_o) : There is no kind of relationship between the VMI measurements allotted to the boundaries, and the weighted choice percentages attached to these boundaries.

Alternative Hypothesis (H_1) : There is a significant positive relationship between the VMI measurements allotted to the boundaries and the weighted choice percentages attached to these boundaries.

After applying the correction factor for ties (T) to both variables, an r_s of 0.63 was obtained. This relationship is highly significant, ($p = < 0.01$) so the Null Hypothesis can be rejected.

4.3.3 Spearman Rank Correlation Coefficient : r_s (3)

(See Appendix 8) For this test, all thirty-four sentences chosen by over 40% of the subjects as boundaries were used.

Null Hypothesis (H_o) : There is no kind of relationship between the number of propositions directly dominated by the proposition at each selected boundary, and the division at which they were selected.

Alternative Hypothesis (H_1) : There is a significant positive relationship between the number of propositions directly dominated by the proposition at each selected boundary, and the divisions at which they were selected.

After applying the correction factor for ties (T) to both variables, an r_s of 0.64 was obtained. This relationship is highly significant ($p = < 0.001$), so the Null Hypothesis can be rejected.

4.3.4 Mann-Whitney U Test (1) (See Appendix 8) In this and the next test bridge sentences or groups are not included unless they have been selected singly by over 40% of the subjects. The extra third division sentence in Passage C (S28) is also omitted. Thus a final total of thirty-two sentences was considered. In none of these tests are the author's own paragraph boundaries included among the extrinsic signals.

Null Hypothesis (H_o) : There is no difference between the percentages of subjects choosing boundaries marked by extrinsic signals, and those choosing boundaries not marked by extrinsic signals.

Alternative Hypothesis (H_1) : The boundaries chosen by the highest percentages of subjects tend to be those boundaries marked by extrinsic signals.

As $U_1 = 47.5$, which is less than 66, this result is highly significant ($p = < 0.01$). Thus the Null Hypothesis can be rejected.

4.3.5 Mann-Whitney U Test (2) (See Appendix 8)

Null Hypothesis (H_o) : There is no difference between the percentages of subjects choosing boundaries marked by intrinsic signals, and those choosing boundaries not marked by intrinsic signals.

Alternative Hypothesis (H_1) : The boundaries chosen by the highest percentages of subjects tend to be those boundaries marked by intrinsic signals.

As n_2 exceeds 20, it is necessary to obtain z. $z = 0.28$, so $p = 0.3897$, which is not significant. The Null Hypothesis therefore stands.

4.3.6 Binomial Tests (1) (See Table 24) In this series of tests and the two following series, an attempt is made to establish the presence of relationships between the various linguistic signals (including VMI measurements of 0.16+) selected by the subjects. No differentiation is made between the divisions at which the choices were made. The sentence added at Division 3 on the basis of the weighted percentage is not included (Passage C, S28). In this particular test, an attempt is made to establish a relationship between subject choices at the different levels, and the number of linguistic signals of any kind present at the same boundaries.

Null Hypothesis (H_o) : $p_1 = p_2 = \frac{1}{2}$, ie there is no difference between the probability of a particular number of signals (3+, 2+, 1+) being present at a boundary chosen at a particular level (Divisions 1-3, 4, and all combined) by the subjects (p_1), and the probability of there being no particular number of signals present there (p_2).

Alternative Hypothesis (H_1) : $p_1 > p_2$, ie the probability of a particular number of signals (3+, 2+, 1+) being present at a boundary chosen at a particular level (Divisions 1-3, 4, and all combined) by the subjects (p_1), is greater than the probability of no particular number of signals being present there (p_2).

Nine Binomial Tests were made on all combinations. As can be seen from Table 24, the following proved highly significant : at least one signal at all divisions ($p < 0.001$ for Divisions 1-3, $p < 0.001$ for Division 4, and $p < 0.00003$

when all four divisions are combined). The presence of at least two signals proved significant at Divisions 1-3 ($p < 0.001$). In these cases, therefore, the Null Hypothesis can be rejected.

In the case of two combinations there is a highly significant negative probability. These are three signals or more at Division 4 ($p = < 0.001$ against), and when all divisions are combined ($p = 0.005$ against).

TABLE 24

(a) Raw Totals

DIVISIONS	TOTAL	3+	2+	1+
1-3	13	6	11	13
4	21	3	8	18
1-4	34	9	19	31

(b) Probabilities

DIVISIONS	TOTAL	3+	2+	1+
1-3	13	$\overline{0.500}$	$\overline{\underline{0.011}}$	$\overline{\underline{<0.001}}$
4	21	$\overline{\underline{<0.001}}$	$\overline{0.192}$	$\overline{\underline{<0.001}}$
1-4	34	$\overline{\underline{0.005}}$	$\overline{0.305}$	$\overline{\underline{<0.00003}}$

NUMBERS OF MARKERS (ALL TYPES) FOUND AT DIFFERENT DIVISIONS

4.3.7 Binomial Tests (2) (See Table 25). An attempt is made to establish a relationship between subject choices at different levels and the three types of linguistic signal present at the same boundary, each type being considered separately. (ie intrinsic, extrinsic and VMI measurements of 0.16+).

Null Hypothesis (H_o) : $p_1 = p_2 = \frac{1}{2}$ ie there is no difference between the probability of a particular category of signal (intrinsic, extrinsic and VMI measurements of 0.16+) being present at a boundary chosen at a particular level by the subjects (Divisions 1-3, 4, and all combined) (p_1), and the probability of no particular category of signal being present there (p_2).
Alternative Hypothesis (H_1) : $p_1 > p_2$, ie the probability of a particular category of signal (intrinsic, extrinsic and VMI measurements of 0.16+) being present at a boundary chosen at a particular level by the subjects (Divisions 1-3, 4, and all combined) (p_1) is greater than the probability of no particular category of signal being present there (p_2).

Nine Binomial Tests were made on all possible combinations. The presence of one or more intrinsic signals with Divisions 1-3 proved significant ($p = 0.046$) - in only this case, however, can the Null Hypothesis be rejected. In the case of two combinations there is a highly significant negative probability : these are the possibility of a VMI movement of 0.16+ being present at a boundary chosen at Division 4 ($p = < 0.001$ against), and its being present where all divisions are combined ($p = 0.005$ against).

TABLE 25

NUMBERS OF INTRINSIC, EXTRINSIC & VMI MARKERS FOUND AT DIFFERENT DIVISIONS

(a) Raw Totals

DIVISIONS	TOTAL	INTR (1+)	EXTR (1+)	VMI
1-3	13	10	9	8
4	21	12	11	2
1-4	34	22	20	10

(b) Probabilities

DIVISIONS	TOTAL	INTR (1+)	EXTR (1+)	VMI
1-3	13	$\overset{+}{0}.046$	$\overset{+}{0}.133$	$\overset{+}{0}.291$
4	21	$\overset{+}{0}.332$	$\overset{+}{0}.500$	$\overset{-}{0}.001$
1-4	34	$\overset{+}{0}.061$	$\overset{+}{0}.195$	$\overset{-}{0}.005$

4.3.8 Binomial Tests (3) (See Table 26) In this series each category of signal is considered separately, and the probability of its coincidence with the subjects' choice of boundary (at any division) is calculated.

Null Hypothesis (H_o) : $p_1 = p_2 = \frac{1}{2}$ ie there is no difference between the probability of a particular signal occurring at a boundary chosen by over 40% of the subjects (p_1), and the probability of its not occurring (p_2).

Alternative Hypothesis (H_1) : $p_1 > p_2$, ie the probability of a particular signal occurring at a boundary chosen by over 40% of the subjects (over 50% in the case of Divisions 1, 2 and 3) (p_1) is greater than the probability of its not occurring (p_2).

The Binomial Test was applied in turn to each of the three intrinsic signals, the four extrinsic signals and VMI measurements of over 0.16. As can be seen in Table 26,

TABLE 26

RELATIONSHIPS BETWEEN SPECIFIC MARKERS AND SUBJECTS' CHOICES

MARKER	TOTAL	TOTAL AS BOUNDARIES	PROB.
Text Reference	13	11	0.011
VM Index 0.16+	12	10	0.019
Short Sentence	13	10	0.046
Contrast Word	6	5	0.109
Lexical Cohesion Chain	21	13	0.192
Particular / General	13	6	0.709
Change of Verb Tense	18	7	0.881

the Null Hypothesis can be rejected in the case of the VMI measurements (with a probability of 0.019), text reference signals (with a probability of 0.011), and short sentences ($p = 0.046$). It should be noted that short sentences at the end of information blocks, ie forming the sentence before as well as after a boundary, were included, as an informal examination of a large number of texts had previously suggested that the short sentence could signal a boundary in both positions.

In all other cases, ie all the intrinsic signals as well as the contrast words, the probability was found not to be significant, so the Null Hypothesis stands.

4.4 General discussion of results Each hypothesis will be taken in turn, and related to the findings of the statistical tests.

4.4.1 Hypothesis 3 (a) The principle behind this hypothesis bears some relationship to Christensen's 'levels of generality', which he applied explicitly to the structure of the sentence, but later carried over by implication to the paragraph (See Chapter 3). The level of generality of a sentence in relation to its neighbours is intuitively real but difficult to formalise. The method used earlier in this chapter (2.2.3) is limited, in that it cannot possibly pinpoint all instances, even of those immediately obvious to the reader.

Meyer found a model based on Grimes' Semantic Grammar of Propositions useful as a means of presenting content structure hierarchically, with superordinate information at the top of the tree and subordinate at the bottom. If it is true that an information block usually begins with the most general content, which is later supported by subordinate content, then this model should also be suitable for the analysis of information chunking. A significant relationship in fact exists in our data between divisions where there is a movement from particular to general and divisions where there is a movement upwards according to the VMI. Of thirteen instances of the former, only three are not accompanied by the latter, this having a probability of < 0.046 according to the Binomial Test.

Table 27 shows the pattern of distribution of VMI measurements over all four divisions. The sentences chosen by over 40% of the subjects are presented with the corresponding vertical movements, however small.

TABLE 27

PATTERN OF DISTRIBUTION OF VMI MEASUREMENTS OVER ALL THREE PASSAGES

	A		B		C	
	S	VMI	S	VMI	S	VMI
1	20	0.50	22	0.63	16	0.09
2	9	0.09	11	0.27	11	0.40
	27	0.27	17	0.16	22	0.08
			24	-		
			34	-		
3	18	0.16			28	-
4	12	0.16	8	0.14	4	-
	14	-	9	0.07	9	-
	16	0.14	13	-	15	-
	17	0.07	25	-	19	-
	29	-	27	0.05	24	0.16
	32	0.07	30	-	31	-
			36	-		
			37	-		
			39	0.12		

The major divisions at 1 and 2, in all cases intended by the subjects to be full paragraph divisions, are, with two exceptions, marked by upward movements. The first exception, S24 in Passage B, is a node dominating a large number of propositions, so will be considered under Hypothesis 3 (b) in the next section. The second, S34, also in Passage B, is the second sentence of a two sentence bridge paragraph (according to the author's own paragraphing). Although this second sentence was the chosen boundary for 55.56% of the subjects, 33.33% chose S33, after a movement of as much as 0.35, the second highest in the passage.

Paragraphs at Division 1 should be the most likely to display large VMI measurements, as here the subjects were given their first choices. In two cases, A and B, the subjects chose the largest vertical movement of the passage, as many of 100% choosing a movement of 0.50 in A, and 72.22% a movement of 0.63 in B. There was far less unanimity in C, perhaps because the sentence chosen by 55.55% of the subjects represented a VMI measurement of only 0.09, by no means the largest in the passage. It did, however, also represent a movement from particular to general.

Passages A and B show a common pattern, with VMI measurements becoming smaller with each succeeding division. The exception is S9 in Passage A. If S8 and S9 (Passage A) are considered together, however, they have as many as five signals, both intrinsic and extrinsic, so are likely to be chosen even without a large vertical movement.

Passage C does not follow the pattern so clearly, although all three sentences chosen at Divisions 1 and 2 do show vertical movements. One major discrepancy is S8, with a VMI measurement of 0.24, not chosen as a division but nevertheless having the second highest VMI measurement of the passage. Passage C in almost every case presented more difficulty to the subjects than did Passages A and B, one reason being its less clear hierarchical structure, particularly at the lower levels.

In spite of these differences, the Spearman Rank Correlation Coefficient results support the hypothesis. For both tests a highly significant relationship was revealed between the vertical distance of the upward movement and the height of the node (as measured by the VMI), and the division at which a paragraph or lower level information block was selected by the subjects. The earlier the choice was made, or the higher the weighted score it received, the more likely was it that the vertical movement would be greater and the node higher. Hypothesis 3a therefore stands.

4.4.2 Hypothesis 3(b) In some cases a division was chosen by over 40% of the subjects where there was no upward movement. In these cases it was hypothesised that the movement across the tree or downwards would depend on the number of propositions dominated by the proposition at the division. Table 28 shows the number of propositions at each division not accompanied by upward movements.

TABLE 28

PATTERN OF DISTRIBUTION OF NUMBER OF PROPOSITIONS AT EACH DIVISION NOT ACCOMPANIED BY AN UPWARD MOVEMENT

PASSAGES

DIVISIONS	A		B		C	
	S	PROPS	S	PROPS	S	PROPS
1		-		-		-
2		-	24	29		-
			34	10		-
3		-		-	28	11
4	14	5	13	7	4	8
	29	21	25	1	9	3
	32	13	36	0	15	2
			37	4	19	3
					31	0

It is immediately obvious, especially if this table is compared with Table 27, that the higher level divisions (1-3) are almost all accompanied by upward movements. S24 and S34 in Passage B are exceptions already mentioned. S24 has only one signal - an intrinsic one - and lies so near the sentence chosen at Division 1 (S22) that it is at first surprising that it should have been selected at all. However, its position at a node dominating 29 propositions through two main branches, one of which has three sub-branches, over-rides any possible discouraging factors. S28 in Passage C is a similar case.

Here the movement is downwards, but although it only dominates eleven propositions, the node has three main branches, with a total of seven sub-branches.

Most of the sentences shown at Division 4 seem to have some other reason for their choice apart from the number of propositions. S29 and possibly S32 (Passage A), however, stand out as exceptions to the pattern. The eighteen propositions directly dominated by the node at S29 represent an extended example, following on from information given in S27 and S28. This could quite well have been selected at Division 3, but many of the subjects are likely to have preferred not to have chosen a paragraph beginning with the words 'for example', an example traditionally being considered as a fairly short specific illustration to a more general statement forming the first part of a paragraph (compare Becker's TRI pattern, Chapter 3). When asked to identify 'smaller segments', however, they had no hesitation in choosing S29, with its eighteen propositions and two branches dividing into six sub-branches. S32, a continuation of the same example, only dominates eleven propositions, but through two branches, one of which has six sub-branches.

The Spearman Rank Correlation Coefficient results indicate clearly that when <u>all</u> sentences chosen by over 40% of the subjects are considered, there is a highly significant relationship between the number of propositions directly dominated by the proposition at the selected boundary, and the divisions at which they were selected.

The earlier the choice is made, the greater the number of propositions which are dominated by the proposition at the selected boundary. Hypothesis 3 (b) therefore stands.

4.4.3 Hypothesis 4 It is postulated here that extrinsic signals which can be manipulated by the writer are better indicators of information block boundaries than are intrinsic signals, which are part and parcel of the content. As has already been noted, the author's own paragraph boundaries are not included among the extrinsic signals, on the grounds that these were not available to the subjects when making their choices.

First, the difference in signal type was ignored, and the number of markers of any type found with paragraph choices at different divisions of the text was considered. It was found that there was a highly significant probability ($p = 0.001$) that two or more signals would occur at Divisions 1-3, and a similar probability ($p = 0.001$) that one or more signals would occur at Division 4. At the same time there was a highly significant negative probability ($p = < 0.001$) against there being three or more signals at Division 4. Thus it seems that the higher in the tree the division is made the more likely it is that a greater number of signals will occur.

Carrying out a similar test, but this time sub-categorising the signals into their three main types (ie intrinsic, extrinsic and VMI measurements of 0.16+), it

was found that the probability of there being an intrinsic signal at a boundary chosen at Divisions 1-3 was just significant ($p = 0.046$). No significance was found for the presence of extrinsic signals, and a highly significant negative probability ($p = < 0.001$) was found against there being a VMI measurement of 0.16+ at a boundary chosen at Division 4. This, of course, does not mean that vertical movements are not found at the lowest choices, as movements there could be (and indeed are) present, but with a measurement of below 0.16.

The above results could, however, be misleading. There are 52 intrinsic as opposed to only 32 extrinsic signals present in these passages, with the result that it is more likely that intrinsic signals will be present, if only through superiority in numbers. It was therefore necessary to look at both types of signal separately.

A glance at Table 29 will show that although intrinsic signals predominate numerically, those boundaries chosen by the highest percentages of subjects are more likely to be marked by extrinsic than by intrinsic signals. Of the three boundaries chosen by a full 100% of the subjects, two are only accompanied by extrinsic signals. According to the Mann-Whitney U Tests, no significant difference was found between the percentages of subjects choosing boundaries marked by intrinsic signals, and those choosing boundaries not so marked, whereas a highly significant difference was found in the case of extrinsic signals.

TABLE 29

DISTRIBUTION OF INTRINSIC AND EXTRINSIC SIGNALS OVER BOUNDARIES CHOSEN BY HIGH PERCENTAGES OF SUBJECTS

	WITH SIGNALS		WITHOUT SIGNALS	
% OF SUBJECTS	60%+	70%+	60%+	70%+
EXTRINSIC	13	11	3	0
INTRINSIC	9	7	7	4

Next, the signals were broken down into their seven types (this time including vertical movements), to see whether there was any difference in the likelihood of each being chosen as a boundary, and binomial tests were used to measure probability. The text reference signals and the presence of a VMI measurement of above 0.16, proved to be the most likely indicators of boundaries. Of a total of 12 vertical movements 10 were chosen by over 40% of the subjects, giving a probability of 0.019, and of a total of thirteen text reference signals, eleven were chosen, with a probability of 0.011. These were closely followed by the short sentence, which was just significant, with a probability of 0.046. The presence of a contrast word was not found to be significant, but as only six instances occurred in these three passages, the number was too small for valid statistical conclusions to be reached. As five of these six were chosen, it seems likely that a larger number of examples would have produced

a statistically significant result.

Of the three types of intrinsic signal, the lexical cohesion chain seems the most reliable boundary marker. Change of verb tense is the signal least likely to indicate a boundary, but the situation might be different if we were concerned with narrative rather than expository prose. In general, the boundaries marked by a high VMI measurement and all three extrinsic signals show a far higher likelihood of being chosen than do those boundaries marked by intrinsic signals.

Hypothesis 4 therefore stands, and the following additional points can be made :

(a) The hypothesis is valid in spite of the fact that intrinsic signals are greater in number than extrinsic signals.

(b) The higher a boundary occurs in the tree structure of a passage, the more likely is it that a greater number of signals will be present.

(c) The extrinsic signals taken separately each show a higher probability of being chosen as boundaries than do the intrinsic signals.

(d) Of the extrinsic signals, text reference is the most reliable marker, and of the intrinsic signals the lexical cohesion chain is the most reliable.

(e) Change of verb tense seems to play little or no part in marking boundaries in expository text of this type.

(f) A high VMI score, which was considered as a separate category, is, second only to the text reference signal the most reliable boundary marker.

5. A comparison with Koen, Becker and Young's experiment (See Chapter 3). Although the method used was similar, Koen, Becker and Young started out with slightly different aims. They wanted, first of all, to find out whether paragraphing was arbitrary or conventional. The present experimenter, however, took for granted the conventionality of paragraphing, and indeed of all text chunking, and as a result required 40% or more of the subjects to choose a boundary before accepting it for analysis. This contrasts with Koen, Becker and Young's lower limit of only 20%.

Other aims of Koen, Becker and Young were concerned with the function of paragraph signals (which they termed 'cues'). They wished to establish that most paragraph signals were formal, that there would be a positive correlation between the percentage of subjects choosing a boundary and the number of signals present, and that there would be a further positive correlation between the percentage of subjects choosing a boundary and the presence of three structural breaks. As we have seen, the present experiment is concerned with text chunking rather than simply with paragraphing, so the subjects were required to divide the text hierarchically, also marking the boundaries of chunks below paragraph level.

The criteria used for identifying signals differ markedly between the two experiments. Becker's theory of paragraph analysis is described in Chapter 3, and in this experiment he modifies his original model to suggest three 'systems' - lexical, grammatical and rhetorical - which

interlock and overlap throughout a text. Thus his signals divide into three groups related to these systems. The present experiment arose from a practical teaching problem, the difficulties second language learners have in writing expository prose, so as a result the divisions chosen are of practical rather than of purely theoretical origin, namely, those signals that are inherent in the semantic content of the text (intrinsic) and those that can be independently manipulated by the writer himself (extrinsic). Upward movements in the tree structure of the text (the VM Index) form a third category, which stands alone.

To a certain extent there is overlap, but both forms of division include categories not present in the other. Koen, Becker and Young do not indicate precisely what their sub-categories are, but it seems that their grammatical system includes grammatical parallelism (including parallel grammatical roles). The categories dealt with in the present experiment and omitted by Koen, Becker and Young are the short sentence, movement from particular to general, and upward vertical movements. The movement from particular to general may be covered by the sub-categories of the rhetorical system, but if this is so the choice of boundary is based on intuitive rather than formal criteria.

Up to a point, the results of the two experiments are similar. Although this is not its primary aim, the present experiment reinforces Koen, Becker and Young's

finding that the paragraph is a conventional unit, and the fact that all the signals used in the analysis were formal suggests that formal markers are at the least very important to the reader in determining information block boundaries. Koen, Becker and Young found the lexical system, represented by the lexical equivalence chain, the least effective signal, and this is partly supported by the present experiment, in that the lexical cohesion chain was less likely to be chosen than any of the extrinsic signals.

No evidence was found, however, that the greater the number of breaks present the higher the percentage of subjects choosing the boundary would be. A high percentage seems more dependent on other factors, like position in the hierarchical structure, or the type of signal present. The highest percentage choices were either in the first and second divisions, and thus probably higher in the tree, or, if the fourth division, accompanied by at least one extrinsic signal, usually text reference. Three or more signals, not necessarily connected with high percentage choices, tended to occur in the first three divisions rather than the fourth division.

CHAPTER 7

THE BRIDGE SENTENCE AND THE TOPIC SENTENCE

1. Introduction

Up to now the investigation has been experimental. In this chapter, however, the conclusions will be more tentative, as they will be based not on specially constructed experiments, but on evidence taken from the experimental passages but incidental to their main purpose. Thus no hypotheses are put forward and little statistical evidence is presented. It is hoped, however, that the conclusions offered will have some validity, and point the way towards future verification. The objects of investigation will be two well-known traditional rhetorical concepts - the bridge sentence and the topic sentence.

2. The bridge sentence

Strictly speaking, a bridge may be any element below, at or above sentence level. In a coherent text there must, of necessity, be a bridge of some kind, implicit or explicit, linking every sentence with whatever precedes it, usually the 'given' or 'known' constituent. A bridge may be of any size, from a single word to a chapter, or even above. For example, Chapter 4 in this work forms a bridge between the survey of the literature (Chapters 2 to 3) and the experimental section (Chapters 5 to 8).

In three places in the passages (A8/9; B24/25; B33/34) the subjects were unable to reach a clear decision as to which of two consecutive sentences constituted the beginning of a second level information block. These areas of indecision have been considered as places in the text where perhaps bridge sentences occur.

2.1 Bridge sentences in the experimental passages Although a bridge sentence can usually be intuitively recognised, a strict definition has been used with respect to those bridge sentences claimed to occur in the experimental passages. A bridge sentence as considered here is defined as a sentence which could equally well be categorised as part of the information block preceding it or as part of that succeeding it. Considering two consecutive sentences, if 40% or more of the total number of subjects choosing the two sentences have chosen each sentence separately as marking a division, then the first sentence is considered to be a bridge sentence. Similarly, if 40% or more of the weighted combined percentages for both sentences together can be allocated to each of the two sentences separately, then the first sentence is considered to be a bridge sentence. This definition should only be applied to sentences selected at Divisions 1 to 3 : at the lowest level, where information blocks are inevitably smaller, consecutive sentences may occasionally be selected because they mark the beginnings of consecutive information blocks. The percentages for the three pairs of sentences appear in Table 30.

TABLE 30

POSSIBLE BRIDGE SENTENCES IN THE EXPERIMENTAL PASSAGES

SENTENCE PAIRS		PERCENTAGE CHOICES	
		DIVISION 2	COMBINED
A8/9	8	41.67	41.38
	9	58.33	58.62
B24/25	24	71.43	61.33
	25	28.57	38.67
B33/34	33	37.50	40.54
	34	62.50	59.46

Using the above criteria, A8 can definitely be considered as a bridge sentence, and B33 can also narrowly be accepted on the percentage of weighted and combined choices. B24 is not accepted, nor is it acceptable for other reasons, which will become clear later in this chapter.

2.2 Rees and Urquhart's examples (1976) In order to investigate the linguistic principles behind the bridge sentence, it is necessary to examine a series of much simpler examples, and for this purpose it is convenient to look at the materials used in an experiment on information chunking conducted by Rees and Urquhart (1976).

Rees and Urquhart were interested in using intonation to elicit their subjects' interpretation of conceptual paragraphs. They selected six short passages, each consisting of two distinct sub-sections linked by what they judged intuitively to be bridge sentences, and asked their subjects first to mark the divisions, and then to read the passages aloud. Their hypothesis was that shift of key would indicate where the subjects were making their divisions, and thus where conceptual paragraphs occurred. The passages appear in full in Figure 23.

FIGURE 23

REES AND URQUHART'S PASSAGES

(The bridge sentences are underlined)

1. Wide publicity has been given during the last few decades to each of the sensational advances of genetics and to their implications for man. Yet the public has had few opportunities to get a comprehensive view of what genetics has accomplished. It was to supply this deficiency that the BBC produced this series on Genetic Engineering. Several of the previous contributers have had occasion to mention that genetic engineering raises conflicts with current standards of conduct. In this final contribution to the series, I have been asked to discuss these conflicts.

2. For its weight, there is really nothing wrong with the mechanical properties of wood. The weight of wooden structures is generally at least comparable to that of metal ones. Moreover, wood is astonishingly resistent to stress concentrations. We pay for this, however, in the vulnerability of wood to moisture. Wood is affected by liquid water in the form of rain, rivers, seas, and so on, with which it may come in contact. More importantly, it is affected by the moisture vapour which is always in the air.

3. Vast tracts of land in the Americas and Asia are yearly becoming more and more barren. In Brazil, for example, thousands of square miles along the Amazon, which until recently supported a rich and economically valuable flora, and fauna, along with a small human population, are now unproductive wastelands. This phenomenon is a direct result of water erosion. In tropical woodlands, the top-soil is highly fertile, but thin and thus vulnerable to rain. It thus requires protection, which in the natural state of its protective covering is liable to be washed away down to barren, unproductive rock.

4. An unactivated nerve fibre maintains a state of chemical stability with concentrations of potassium inside and outside the lining membrane in a ratio of 30:1. The end of a nerve fibre is not structurally joined to the next cell, but the small gap between them can be bridged chemically. This functional junction is known as a synapse. Not all the chemicals which are transmitters of the synapse are known. Among the most important, however, are acetyl choline and noradrenaline.

5. In the heart of desert regions the distinctive qualities of desert climate are unmistakable. Drought, sun, wind, occasional rain storms, and heat by day characterise the climatic year. Things are different to this, however, on the desert margins. Unless a desert region is bounded by highland, desert climate and desert scenery vanish by replacement. In such circumstances the status of a given locality may be very doubtful, and the limits of the true desert very difficult to fix.

6. The subject of innate equality has become confused with the political question of equality of opportunity. The main source of the confusion has been the attempt by some to claim especial virtues for particular classes. This claim has been especially associated with eugenic propaganda. The founder of the eugenic movement was an English-man, Francis Galton. Galton was a pioneer of the use of statistical methods in biology. He became the first President of the Eugenics Education Society, founded in London in 1908.

At first sight it seems that Rees and Urquhart judged correctly in their choice of bridge sentences, as, of 66 responses (6 times 11), 31 began their second paragraph with a 'bridge sentence' and 27 ended their first

paragraph with one. (Eight chose other divisions). On looking at the individual passages separately, however, there is marked variation between the responses, and if the same criterion is used as was used in the previous section, only two of the six appear to be true bridge sentences (See Table 31).

TABLE 31

BRIDGE SENTENCES : RESULTS OF REES AND URQUHART'S EXPERIMENT

	SUBJECTS' DIVISIONS		
PASSAGES	a	b	c
1	6	5	0
2	6	2	3
3	5	5	1
4	2	5	4
5	10	1	-
6	2	9	-
TOTAL	31	27	8

KEY

a = 'bridge sentence' at beginning of second paragraph

b = 'bridge sentence' at end of first paragraph

c = division placed elsewhere

From these results it appears that only passages 1 and 3 contain true bridge sentences. Of the rest, those in 5 and 6 are definitely not bridge sentences, while those in 2 and 4 are indeterminate, so many subjects having made other choices.

This conclusion is surprising, bearing in mind the similar structure of all six supposed bridge sentences. Each one consists of the following features :

1. At the <u>beginning</u> of the sentence:

   Either (a) The demonstrative pronoun 'this' with text reference back to the immediately preceding information block

   Or (b) The demonstrative adjective 'this' followed by a general noun, (eg. <u>statement</u>, <u>suggestion</u>, <u>phenomenon</u>, <u>situation</u>) with text reference back to the immediately preceding information block.

2. At the <u>end</u> of the sentence :

   A noun phrase at a similar or somewhat less general level which states the main topic of the succeeding information block or blocks. Thus it may act as a topic sentence, but it is more likely that succeeding information blocks contain their own more fully developed subordinate topic sentences at a lower level.

3. In the <u>middle</u> of the sentence :

   Often a word or phrase expressing a relationship of contrast or result is present. Examples from Urquhart

and Rees's passages are however (2, 5) result (3) and different (5) (Compare Kane and Peters' adversative and cause and effect categories (Kane and Peters, 1966 : 105-113))

On closer examination, however, two quite different factors appear to be vital to the identification of the bridge sentence, these being the fulfillment of the reader's expectations, and the distribution of 'given' and 'new' information.

2.2.1 The fulfillment of the reader's expectations It seems essential that the second part of the sentence should state the topic of whatever succeeds it. If expectations are then not fulfilled, the reader will tend to place the sentence at the end of the previous information block.

This is convincingly illustrated in Passage 6 (Figure 23). Here the supposed bridge sentence promises that the subject of the next block will be 'eugenic propaganda', but instead of this a brief description of Francis Galton follows. It is true that Galton was the first president of the Eugenics Education Society, but the development of 'propaganda' is never actually reached in this short extract. In Passage 4 there is a similar deficiency. The reader is expecting information about the synapse, but what he receives is limited to the chemicals that transmit it. A tendency towards the same choice as in Passage 6 occurs here, but it is not quite so marked, as more subjects make other choices. Both sentences

might function more obviously as bridges if their texts were continued further.

2.2.2 The distribution of 'given' and 'new' information

A more important factor appears to be the distribution of 'given' and 'new' information in relation to theme and rheme. Before discussing this in detail, however, it is necessary to say something about the relationship of these two simultaneous text structures, noted by M.A.K. Halliday. Halliday (1967) distinguishes 'information focus', ie the distribution of given and new information, from 'thematisation', ie the presence of theme and rheme. The former is concerned with the organisation of the information structure of the sentence, while the latter indicates where the message lies.

One major difference between the two is their relative flexibility. In the words of Hausenblas (1969), quoted by Daneš (1974), the theme is 'what has been posited to the fore, into the focus of the field of vision, and, at the same time, what presents a foundation to be developed (elaborated) in the subsequent discourse'. Thus the theme never moves, unlike 'given', which in its unmarked form stands in first position, but does not necessarily always do so. The normal situation is for 'theme' and 'given' to coincide in first position, but occasionally, in the marked cases, theme will coincide with 'new'. The difference can best be summed up by Halliday :

> Basically, the theme is what comes first in the clause; and while this means that . . . . there is in the unmarked case (ie if the information structure is unmarked) an association of the theme with the given, the two are independent options. The difference can perhaps best be summarised by the observation that, while 'given' means 'what you were talking about' (or 'what I was talking about before') theme means 'what I am talking about' (or 'what I am talking about now') ; and, as any student of rhetoric knows, the two do not necessarily coincide.
>
> (Halliday 1967 : 212)

In texts in general there appears to be a strong tendency for information blocks to begin when 'new' and 'theme' coincide. A bridge sentence should therefore follow the unmarked pattern of 'theme' corresponding with 'given', and it should be succeeded by a sentence of the same pattern. If the succeeding sentence reverses the pattern, then there will be a strong tendency for subjects to choose that sentence as the beginning of a new information block.

2.2.3 Application to Rees and Urquhart's examples The relationship between the distribution of 'theme', 'given' and 'new' in the sentences under consideration are shown in Figure 24.

FIGURE 24

ANALYSIS OF REES AND URQUHART'S BRIDGE SENTENCES

1. It was [to supply this deficiency] that the BBC produced this series on genetic engineering

2. [We] pay for this, however, in the vulnerability of wood to moisture.

3. [This phenomenon] is a direct result of water erosion

4. [This functional junction] is known as a synapse

5. [Things] are different to this, however, on the desert margins.

6. [This claim] has been especially associated with eugenic propaganda.

KEY

[ ] = theme

——— = new

----- = given

In four sentences out of the six (1, 3, 4, 6) theme and given coincide. In the remaining two sentences (2 and 5) the theme contains completely new information. This suggests that these two sentences, 2 and 5, cannot be

classified as bridges. Table 32 summarises how the fulfillment of expectation and the distribution of given and new information taken together act as indicators of bridge sentences. The status of the sentence is left undefined when the subjects did not limit their choices to the sentences in question.

TABLE 32

ANALYSIS OF REES AND URQUHART'S BRIDGE SENTENCES

SENTENCE NO	SUBJECT CHOICE	FULFILLMENT OF EXPECT.	GIVEN = THEME	RESULTING STATUS
1	6/5	Yes	Yes	Bridge
2	6/2(3)	Yes	No	Doubtful
3	5/5(1)	Yes	Yes	Bridge
4	2/5(4)	No	Yes	Doubtful
5	10/1	Yes	No	Not bridge
6	2/9	No	Yes	Not bridge

According to this analysis, the only sentences which can truly be considered as bridge sentences are 1 and 3, while 5 and 6 must be rejected.

2.3 Application to the experimental passages Rees and Urquhart do not give their sources, but it is unlikely that these examples are truly authentic. Indeed, reference to the original version of Biddulph's passage would show that Sentence 5 has in fact been 'doctored' for the purpose of the experiment. This is, of course,

legimate, but we should not expect to find quite such ideal examples when we turn to authentic text.

The three possible bridge sentences in the experimental passages will now be examined, to find out how far they correspond to the guide-lines laid down above. Each will be examined within its immediate context.

2.3.1 Passage A, Sentence 8 (Figure 25)

FIGURE 25

PASSAGE A, SENTENCE 8 : ANALYSIS

A7 [In the arid regions of the world] the wind is a powerful force in removing material from one area to another

A8 [All this] is natural

A9 [But Nature] has also provided certain defensive forces

A10 [Bare rock surfaces] are in due course protected by soil . . .

KEY

[ ] = theme

——— = new

- - - - = given

Lines 8 and 9 in this extract pose several interesting problems. The author himself did not make a

paragraph here, yet a total of 63.16% of the subjects placed a Division 2 boundary at either A8 or A9. It is clear, however, that the author intended a transition, if not an actual paragraph boundary, marking it as he did by two extrinsic and three intrinsic signals.

At first glance the short sentence (A8) seems meaningless, or at least without purpose, but on closer examination it becomes clear that if A8 and A9 were coordinated they would make together a bridge sentence exactly of the type present in Rees and Urquhart's example, with the three features described in the previous section (2.2). The subjects were therefore placed in a quandary. Either they should interpret A8 and A9 as two separate sentences, in which case A8 would belong to the first paragraph and A9 to the second, or they should interpret them both together as a bridge sentence, in which case they should both belong as a unit to one paragraph or the other. It seems that the fact that A8 is an exceptionally short 'short' sentence militated against A9 being placed at the end of the first paragraph. Marginally more of the subjects (36.84%) chose to separate the two sentences into separate paragraphs than to put them together at the beginning of the second paragraph (26.32%).

A10 poses a slight problem. At first sight the theme ('bare rock surfaces') seems to be new, which should militate against A8 and A9 being placed before it in the same paragraph. The question arises here whether the

given element in a sentence refers back to the previous sentence only, and if not, how far back reference can legitimately be allowed before the sentence element is labelled 'new' rather than 'given'. Chafe's viewpoint is helpful here. He prefers to look at the situation from the psychological point of view, considering 'given' information as information 'assumed to be in the addressee's consciousness' (Chafe 1974 : 112), and 'new' information as the opposite. An interesting question here is the duration of givenness, the human mind only having limited capacity, and Chafe suggests that the speaker usually considers a given element to have left the addressee's consciousness after at least one sentence has intervened.

Chafe was thinking primarily of the spoken language ; written language being less transitory, it is reasonable to suppose that a greater gap than one sentence is acceptable. Looking again at A10, 'bare rock surfaces' has already occurred in A6, but if this is considered too far back it is also implicit throughout the first part of the passage, which describes the erosive process. Thus, 'bare rock surfaces' can be classified as 'given' information, and as 'given' corresponds with 'theme' there is no overwhelming pressure to make A10 the beginning of a new paragraph.

Thus to a certain extent A8 and A9 follow the pattern suggested above. It would be revealing to find out what the reactions of the subjects would be if the two sentences were coordinated into one bridge sentence.

2.3.2 Passage B, Sentence 33 (Figure 26)

FIGURE 26

PASSAGE B, SENTENCE 33 : ANALYSIS

B32 [This measure of integration] can be supplemented by other measures, based on the market area, newspaper subscriptions, retail trade, public transport, and the like.

B33 [Despite their apparent comprehensiveness,] the SMSA definitions have still not solved the problem of urban boundaries.

B34 [An improved definition using county blocks and commuting data] has been suggested by a team of Chicago geographers.

B35 [Their recommendations] were threefold.

KEY

[ ] = theme

——— = new

- - - - - = given

Although B33 and B34 are longer than A8 and A9, a similar interpretation can be placed on them.

B33 is almost entirely 'given' information. 'Despite their apparent comprehensiveness' refers directly back to the

preceding nine sentences, but 'the problem of urban boundaries', at least when considered in isolation, could also hardly be considered as new, expressing as it does the overall topic of the passage. B34, on the other hand, is almost entirely new information, though the fact that a 'definition' is being introduced, and that it is 'improved', links this sentence with previous ones. Again, if these two sentences were joined, they would form a bridge sentence closer to the model, if somewhat long. The implicit relationship is one of result, and so would therefore be a suitable coordinator.

The subjects, therefore, had two possible courses of action, as in the previous example. Either they should consider the two sentences separately, putting B33 into the first paragraph and B34 into the second, or they should consider them together as a composite bridge, and place them at the end of the first or the beginning of the second. In fact, 55.56% of subjects took the first course of action and 33.33% the second, almost exactly the same proportion as with the first example.

There is no doubt that the author intended these two sentences to together form a bridge, as he chose to isolate them from their surrounding context in a short transitional paragraph, again a point in the experimental passages where the author's original paragraphing and that selected by the subjects did not agree. It is not clear why none of the subjects suggested a boundary at B35,

especially as this, one of the author's original boundaries, is a short sentence. Perhaps the pronoun their, which reduces the newness of the theme, is a contributing factor.

2.3.3 Passage B, Sentence 24 (Figure 27)

FIGURE 27

PASSAGE B, SENTENCE 24 : ANALYSIS

B23 [One definition of world metropolitan areas by demographer Kingsley Davis] runs to twelve pages, including two pages on difficult cases.

B24 [In the United States] the concept of the SMSA was introduced in 1960, so that metropolitan areas could be defined realistically by using three criteria.

B25 [First a population criterion :] each SMSA must include one central city with 50,000 or more inhabitants.

B26 [Special rules] allow contiguous cities and nearby cities to be combined.

KEY

[ ] = theme

——— = new

- - - - - = given

55.56% of subjects chose B24 as a boundary, as opposed to 22.22% choosing B25, which does not put it

within the 40% limit decided on above. In fact, B24 is clearly not a bridge sentence, but the second specific example of the type of definition mentioned in B22.

B24 seems the obvious place to begin a new information block, as it begins with a correspondence between theme and new and the reader's expectations are explicitly expressed and immediately fulfilled. Some subjects, however, seem to have been influenced by the signal first and the partly new theme to make their boundary break one sentence later. As has been shown in Chapter 6, extrinsic reference signals have a strong positive effect on the reader.

Either boundary can be interpreted as legitimate. A boundary at B24 would single out the second specific instance together with its supporting description. A boundary at B25 would leave the two instances together as supports for B22, and isolate the long and rather detailed description of the second. On the whole, the subjects preferred the first alternative, though probably would not have done so if the relationship between B23 and B24 had been more clearly indicated, by means of, for example, parallel grammatical structuring.

2.4 Bridge sentences in general Although bridge sentences are a common feature in expository writing, the relevant literature has more often than not ignored them. Even in the area of traditional rhetoric, of the fifteen texts considered in detail in Chapter 2 only Kane and Peters (1966)

deals with this topic at any length. One problem is that a 'bridge' inevitably contains two elements within one unit, and this does not easily fall within the segmental type of analysis favoured by most linguists, and even by rhetoricians.

Although the Tagmemicists have not directly tackled the problem, their wave view of language offers the most promising frame within which to consider the bridge sentence. Pike and Pike point out :

> (The wave perspective) demands identification of nuclei as peaks of a grammatical wave for its prime point of attention. At the border of two units identified by nuclei, the meeting-place may be fused so that one phonological or grammatical or referential unit may serve simultaneously as the end margin of one unit and the beginning margin of the next one (ie the two units are fused with the joint margin serving as a transition between them). In these instances, the total two-unit complex can be listed, without distortion, as a single entity.
>
> (Pike and Pike 1977 : 27)

As constituent units become smaller their boundaries become more restricted and thus more easily defined. The wave view of text organisation is therefore particularly valid at super-sentential level.

3 The topic sentence

Although some nowadays would deny its validity, the topic sentence has been stressed by all traditional teachers of rhetoric from Alexander Bain onwards as the essential anchor of, and point of entry into, the expository paragraph.

3.1 Some possible definitions The topic sentence can be considered from four possible points of view. It can be defined as

(a) the sentence which states the topic of the paragraph, or indicates its scope. This is the traditional pedagogic definition, and according to this view the topic sentence is optional. Thus its identification is subjective, and often questionable.

(b) the most general sentence in the paragraph This definition is implicit in Becker's (1965) use of the term, and also results in subjective identification procedures.

(c) the sentence upon which all the other sentences in the paragraph depend This is Christensen's (1965) definition, and it follows from his view of the paragraph as a 'macro-sentence', with the topic sentence analogous to the main clause. The main difficulty here is that he does not make explicit the precise ways in which the following sentences depend on the topic sentence, so that the identification of the topic sentence itself is subjective.

(d) the first sentence of the paragraph This definition treats the topic sentence as the theme (in Halliday's terms) of the paragraph, ie its taking-off point, and as every paragraph must of necessity have a first sentence, there are no problems of identification. However, one result is that the pedagogic usefulness of the topic

sentence concept is lost. It seems more productive, therefore, to call the first sentence the 'theme sentence'. This would usually, but not necessarily, coincide with the topic sentence, which would then need to be redefined.

3.2 An alternative definition A major problem of these definitions is that they are all sentence based. While this is understandable, bearing in mind that the main interest of all the workers in this field has been pedagogic, it ignores the fact that related topics within a text are organised hierarchically, and are manifested at different levels above and below the sentence. Thus, according to Jones, who uses the term 'theme' to represent what is conceived here as 'topic' :

> Low level structures such as clusters of two or three sentences within a paragraph have low level themes. Higher level structures, such as a cluster of five or six or so sentences, have higher level themes. The highest level structure has the highest level theme . . . . There is a taxonomic hierarchy of themes within themes. The lower level themes occur as more specific details within higher level themes. The hierarchical structure of themes within a text is a relationship of more specific to more general as one moves up the hierarchy.
>
> (Jones 1977 : 37)

Thus the concept of the topic proposition, represented at surface level as a clause, may be easier to pin down than the more traditional concept of the topic sentence. The following operational definition is suggested :

A topic proposition is any proposition which directly dominates at least three other propositions.

A topic sentence is any sentence which contains no more than three propositions, and whose main proposition is a topic proposition.

(The term 'proposition' here refers to the proposition as defined by Grimes (1975), and as used in the experimental work of the previous chapter).

This definition has several advantages. Although it makes no direct appeal to meaning, it incorporates at least one of the significant elements of more content-bound definitions. For example, if the topic sentence dominates a node, it is likely to be more general than the nodes beneath it, in that the higher levels of a tree are more general than the lower levels. Also, the sentences beneath the topic sentence in the tree must, inevitably depend on it.

The definition is operational in that it supplies a cut-off point for the identification of topic sentences in actual texts, but thus has the disadvantage that it does not provide for lower-level topic sentences and parts of sentences which dominate too short a portion of text to fall within the definition. It could, indeed, be claimed that the function served by the topic sentence in the paragraph is similar to that served by the main clause in the sentence, this being the stand taken by some of the very early prescriptive writers such as Bain and Genung, as well as Christensen. However, the more familiar view of the topic sentence is that it is a statement at a fairly high level of generality, so such lower level

manifestations will be left out of consideration here.

The picture which emerges is nevertheless one of topic sentences at various levels, those at the lowest levels dominating chunks of text subordinate to and contained within larger chunks. The first sentence is frequently the overall topic sentence of the whole text (called by traditional rhetoricians the 'theme sentence'), its subordinate topic sentences being those most likely to be chosen by the author as the topic sentences of his paragraphs. Orthographic paragraphing is not, however, a necessary element, as beneath these major topic sentences are ranged successive layers of minor ones (traditionally termed sub-topic sentences), dominating chunks often too small to be isolated as single paragraphs in their own right. This pattern can be illustrated from the three experimental passages. Their topic sentences according to the above definition are identified in Figure 28.

It should not be forgotten that as the expression of the topic may occur at below sentence level it may also occur above sentence level, in the form of a higher level information block such as an introductory paragraph or chapter.

3.3 Topic sentences as macro-propositions If topic sentences are manifestations of the macro-structure and represent explicit macro-propositions, as claimed by Van Dijk (1977), then a succession of topic sentences should present a general outline of the content of the text. This is the case with

FIGURE 28a

TOPIC SENTENCES

PASSAGE A

EVERY GEOLOGIST IS FAMILIAR WITH THE EROSION CYCLE

No sooner has an area of land been raised above sea level than it becomes subject to the erosive forces of nature.

The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets and then into rivers and out to sea. The frost freezes the rain water into cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust. Blocks of rock dislodged at high levels are brought down by the force of gravity. Alternate heating and cooling of bare rock surfaces causes their disintegration. In the arid regions of the world the wind is a powerful force in removing material from one area to another. All this is natural.

But Nature has also provided certain defensive forces.

Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks.

Slowly but surely, different types of soil, with differing profiles, evolve, the main types depending primarily on the climate.

The protective soil covering, once it is formed, is held together by the vegetation.

Grass and herbaceous plants, with long, branching, tenuous roots, hold firmly together the surface particles.

The same is true with the forest cover.

The heaviest tropical downpours, beating on the leaves of the giant trees, reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil.

The soil, thus protected by grass, herbs and trees, furnishes a quiet habitat for a myriad varied organisms :

earthworms that importantly modify the soil, bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants.

Chemical action is constantly taking place : soil acids attack mineral particles, and salts in solution move from one layer in the soil to another.

We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions and the parent rock substances develops.

It is a long process, and in many areas there has not been sufficient time, in the geological sense, for completion of the process : the soils are immature.

NOW LET US CONSIDER WHAT HAPPENS WHEN MAN THE PIONEER COMES ALONG

He ploughs up the natural grasslands.

He removes the various branching roots that have held together the surface particles, so that they are now easily moved by the action of rain and wind. In ploughing he has mixed together the surface layers, and provided his crops with a medium in which they can grow and develop, and in which natural plant food is present in varying degrees, or which can be supplemented by animal or chemical manures.

But he has also exposed the soil to the action of the atmosphere.

In some parts of the world exposure is useful : in a cold climate the breaking up of the clods of clay by frost action is beneficial. At the other extreme, as we have already pointed out, in tropical climates, by exposure to the atmosphere such rapid chemical action is set up that the natural plant food in the soil is quickly destroyed.

Broadly speaking, the position is worse when the natural vegetation is forest.

Not only does man the pioneer rob the soil of its source of humus, the fallen leaves, but he exposes a soil quite unused to the direct rays of the sun and the direct fall of the rain to the immediate influence of both.

For example, much of the upland and west coast of Scotland was once forested with the beautiful Scots fir.

The heavy rain, falling on the close pine woods, trickled gradually to the ground and soaked into the soil. Much was evaporated from the leaves, and the floor of the forest, covered with pine leaves, remained comparatively dry, supporting a sparse cover, or undergrowth, of various shade-loving plants or low shrubs, such as bilberry or heather.

When the forests were cut down the heavy rain fell straight on to the surface soil more rapidly than it could drain away.

Especially where there was no steep slope the water was held up, and moisture-loving plants began to flourish, particularly sphagnum, or bog moss. Once the sphagnum was established it acted as a sponge. True, it prevented soil erosion, but it grew and grew until great thicknesses of moss blanketed the whole countryside.

Thus huge stretches of boogland, that known to the botanists as 'blanket bog', extending over wide areas of Scotland and Ireland, are directly due to man's action.

LET US CONSIDER THE VARIOUS DEFINITIONS OF URBAN SETTLEMENTS.

The definitions used in legal and administrative documents will tell us precisely what we mean by Melbourne Australia or Topeka Kansas.

Unfortunately the legal and administrative borders of cities are often a historical or constitutional legacy.

Typically, the legal city has fixed boundaries, that survive long after urban development has exceeded these bounds.

Thus the legal city is often underbounded.

Parts of the urban area may remain outside the city, but share a common boundary with it. Beverley Hills, completely surrounded by the city of Los Angeles, is a case in point.

In England, where administrative regions continue to have a strong administrative quality, some boroughs still have a municipal status that is a legacy of their former importance, and is out of line with their present small size.

The number of inhabitants an area must have to be considered urban also varies from country to country.

In Iceland, places with a few hundred people are termed urban, whereas in the Netherlands a population of 20,000 is needed.

A second approach to defining urban settlements is to ignore the legal boundaries and try to define each settlement in terms of its physical structure.

For example, we might define a settlement on the basis of a continuous distribution of housing, or population above a certain density, or the intensity of traffic.

But there are difficulties here too.

What do we mean by 'continuous housing', and what happens when different definitions don't all give the same answer?

Figure 14.2 presents some different definitions of New York, based on both its legal boundaries and its physical structure.

Note that New York City itself is only a small part of the continuous urban sprawl that is Greater New York.

The mismatch between the legal and the physical city becomes vitally important when the legal city, with its static or declining population, and limited tax base, has to provide public services, like transport or police, for the millions of commuters who cross its boundaries for work each day.

As the discrepancy between the legal and economic boundaries of the city becomes worse, the pressure for some form of revenue-sharing or boundary adjustment grows.

This discrepancy also affects our ability to answer even the simplest questions about the size of the city.

To take an extreme case, the 'legal city' of Sydney, Australia, in 1955 had a population of only 193,000, while the 'built-up' area of Sydney had a population of 1,969,000.

This difference of over nine times in size is unusual, but important enough to make the definition of settlement a matter of concern.

As a result of this problem, international and indeed intranational definitions of urban settlements are being standardised.

One definition of world metropolitan areas by demographer Kingsley Davis runs to twelve pages, including two pages on difficult cases.

In the United States the concept of a Standard Metropolitan Statistical Area (SMSA) was introduced in 1960, so that metropolitan areas could be defined realistically by using three criteria.

First, a population criterion : each SMSA must include one central city with 50,000 or more inhabitants. Special rules allow contiguous cities (ie those directly adjoining each other) and nearby cities (within twenty miles of each other) to be combined.

Second, the metropolitan character of an area is taken into account. At least 75% of the labour force of the country must be employed by non-agricultural industries.

Other criteria for SMSAs relate to population density, the contiguity of townships, and ratios between the non-agricultural labour forces making up the unit.

Finally, the integration of the areas that constitute the SMSA is considered.

Counties are integrated within the county containing a central city if 15% of the workers in the county live in the city. This measure of integration can be supplemented by other measures based on the market area, newspaper subscriptions, retail trade, public transport and the like.

Despite their apparent comprehensiveness, the SMSA definitions have still not solved the problem of urban boundaries.

An improved definition using county blocks and commuting data has been suggested by a team of Chicago geographers.

Their recommendations were threefold.

First, counties of equivalent units were to be retained as the basic building blocks of any system.

Second, counties were to be classified into functional economic areas (FEAs) on the basis of county to county commuting data.

An FEA would consist of all the counties in which the proportion of resident workers who commuted to a given central county (usually containing a city of 50,000 or more inhabitants) exceeded the proportion who commuted to alternative central counties.

Third, FEAs were to be grouped into consolidated urban regions (CURs) when two or more FEAs sent at least 5% of the workers in the central county of one commuting area to the central county of others.

FIGURE 28b TOPIC SENTENCES

PASSAGE B

IN THE FOREST COMMUNITY EVERY INDIVIDUAL STRIVES TO OBTAIN FOR ITSELF AND ITS OFFSPRING THE SPACE AND THE NOURISHMENT NECESSARY FOR LIFE.

This involves competition.

The struggle for existence, which plays such an important part in nature, is exemplified in a most striking manner in the forest.

Above ground it resolves itself mainly into a struggle for light.

Every tree competes with its neighbour for as large a share as possibleof the light falling upon the forest from the sky, by striving to thrust its crown, with its assimilating organs the leaves, above those of its neighbours.

The individuals which fail to hold their own in the struggle are doomed to destruction sooner or later.

Their crowns are overshaded, their assimilation drops, and, left behind in the semi-darkness below, they ultimately fail to nourish themselves sufficiently to maintain life.

Owing to the widespread roots of forest trees, competition also goes on in the soil for water and soil nutrients.

The smaller plants, bushes, herbs, mosses, etc., are also involved in competition.

While they have, as a rule, to submit to the dominance of the trees, they compete amongst themselves and with the younger progeny of the trees for such light as can reach them, and for their share of the moisture and plant food in the soil.

IN A TRUE COMMUNITY, HOWEVER, BESIDES COMPETITION THERE IS ALSO COOPERATION AMONG THE MEMBERS.

The trees protect and support each other against the wind, and their crowns unite to form a canopy which shades the soil, and so keeps injurious weeds out of the forest. They protect, by the shelter they give, not only their own but their neighbours' offspring from destructive frosts. They produce within the forest a climate quite different from that outside. They also produce distinctive soil conditions, which are almost essential for many members of the community, and are, for others at any rate, very beneficial.

The long-continued interaction of competition and cooperation tends to produce ultimately a condition of equilibrium in the forest community which gives it stability and permanence.

When this condition is reached, the forest is said to be the climax type of forest for that particular locality.

The species and their relations to one another best adapted to the conditions of the locality have been established. The climax type is the end of a succession of types of vegetation which may have occupied the site in the past, each of which, by modifying the climate and soil of the site, has paved the way for its supercession by another type. New species have found conditions favourable for establishing themselves, and more or less completely ousted the old. The climax type depends on the climate and soil, and the particular species of plants and animals which have had the opportunity of entering the community, and their actions and reactions on each other.

The equilibrium established in the forest community is not static but dynamic.

It constantly undergoes fluctuations through variations in the weather from year to year, which favour or discourage certain members or sets of members.

There are, however, always compensating factors.

Wet years are succeeded by dry years, warm seasons by cold ones. If one set of members multiplies unduly, its enemies also increase, and the over-produced species is brought back to normal.

The mechanism of the biotic environment is so constructed that it always re-establishes itself, provided there are no overpowerful or persistent interferences from outside.

The most important of these interferences are those for which man is responsible.

He has introduced other members into the forest community, such as new species of trees, grazing animals, etc., and, with axe and fire, has introduced new factors into the environment, favouring some and injuring or destroying other sets of members, with varying effects on the forest. Some of these effects are direct and evident, some indirect and not immediately obvious, but nevertheless of far-reaching consequences. It is therefore necessary that silviculture should be based, not only on a knowledge of the natural relationships of the forest to its environment, but also on the responses of the forest to human action of various kinds.

FIGURE 28c TOPIC SENTENCES

PASSAGE C

all the topic sentences identified in the three passages. If successive topic sentences are read without their supporting text, it will be seen that each resultant passage represents a coherent outline of the main theme (See Figure 28). Thus Jones (1977 : 49) suggests that a topic sentence is always a 'minimum generalisation', or a 'higher level paraphrase'. Such an approach could be of value in the formalisation of a theory of outlining and summary writing.

3.4 The length of the topic sentence The topic sentence of a paragraph tends to be relatively short. According to the definition suggested above, it should not normally contain more than three propositions. When the sentence lengths of seventy representative paragraphs from each of ten regional geography textbooks were analysed, it was found that for eight of these ten the majority of first sentences had below average sentence lengths for their paragraphs (See Table 33). These results, however, were only statistically significant for three of the eight books. Of course, the first sentence of a paragraph, the theme sentence, need not also be a topic sentence.

TABLE 33

ANALYSIS OF THEME SENTENCE LENGTHS IN TEN SELECTED TEXTBOOKS

(A full list of the books will be found in Appendix 9)

	PERCENTAGES		
	-	+	
1. PATTERSON	70.58	29.41	Majority of theme sentences below average sentence length - statistically significant at level of 0.05 (Binomial Test)
2. HOOSON	69.13	30.86	
3. SINNHUBER	63.41	36.58	
4. MORGAN	55.88	44.11	Majority of theme sentences below average length, but not statistically significant
5. HOUSTON	54.28	45.71	
6. WALKER	52.23	49.25	
7. SHORTER	51.42	48.57	
8. WATSON	51.35	48.64	
9. TREGEAR	43.66	56.33	Majority of theme sentences above average length
10. MELLOR	43.47	56.52	

KEY

- = Theme sentences below average length of total sentences in paragraph

+ = Theme sentences average or above average length of total sentences in paragraph

Considering information blocks rather than paragraphs, short sentences may represent sub-topics within the orthographic paragraph rather than at its beginning. Indeed, it was suggested earlier that this is the normal

position for the short sentence (Chapter 5, 2.2.2). Change of topic at the beginning of the paragraph is already strongly marked by indentation, and so a further signal may not be necessary. Indeed, it should not be overlooked that instead of using the traditional short incisive topic sentence, an author may present his theme in only part of a sentence, supporting and elaborating it within the same sentence by means of subordinate clauses (eg Passage C, Sentence 19).

The use of the short sentence to introduce a new topic within the orthographic paragraph is illustrated in Table 34. Of a total of eleven paragraphs (marked as such by the authors) in the three experimental passages, only three begin with short sentences (as defined in Chapter 6, 2.1.2). Of a total of thirteen short sentences, a majority of eleven are designated as topic sentences according to the definition of the topic sentence presented in this chapter. Thus a large majority of the short sentences act as topic sentences for lower level information blocks within the author's orthographic paragraphs.

TABLE 34.

THE FUNCTIONS OF THE SHORT SENTENCE IN THE EXPERIMENTAL PASSAGES

	PASSAGE A			PASSAGE B			PASSAGE C		
	a	b	c	a	b	c	a	b	c
	8	-	-	13	-	✓	2	-	-
	14	-	✓	35	✓	✓	4	-	✓
	21	-	✓				9	-	✓
	24	-	✓				11	✓	✓
	27	✓	✓				24	-	✓
							28	-	✓
TOTALS	5	1	4	2	1	2	6	1	5

a = Short sentences (sentence numbers)

b = Short sentences chosen by the authors as theme sentences

c = Short sentences functioning as topic sentences

The designation of both 'short sentence' and 'topic sentence' is arbitrary, and depends on the assumption that a discrete cut-off point is possible. In reality, both occur on a cline, and any definition must be unsatisfactory in the transitional zone. Like the bridge sentence, this is an area of language best viewed from the wave rather than the particle perspective (Pike, 1959).

CHAPTER 8

PARAGRAPH LENGTH

1. Introduction

It has already been pointed out in Chapter 2 that some writers on traditional prescriptive rhetoric have suggested that paragraph length may vary according to the type of format used (eg Hulbert and Hulbert (1929), Kane and Peters (1966)). In other words, a passage written, for example, in double columns will probably be divided into more paragraphs than one written in a single column across the page with narrow margins. It is common knowledge that 'journalistic' paragraphs tend to be short, even sometimes of one sentence only, though how far this is imposed by the format and how far by the necessity for quick and easy comprehension is debatable.

If it could be shown that the paragraphing of the same passage may vary according to length of line or spacing of print, then the traditional concept of the unified organic paragraph would no longer be acceptable. The paragraph division could then be seen simply as an orthographic device for indicating the hierarchical organisation of the text, by singling out or grouping together one or more information blocks dominated by a common node in the tree. The number of blocks per paragraph would depend on external factors, associated with

aesthetics and comprehensibility rather than on anything inherent in the paragraph itself.

2. D.S. Knapp's experiment At least one experiment on these lines has already been undertaken. D.S. Knapp (1967) presented 394 graduate subjects with an unindented abstract from Chapter 9 of Albert Marckwardt's 'American English', and required them to insert their own paragraph divisions. Sixty-five subjects returned Format A (two columns, élite type, single spaced), 58 returned Format B (narrow margins, élite type, single spaced), 194 Format C (single spaced, pica type), and 77 Format D (double spaced, pica type). Knapp expected Format A to receive the greatest number of divisions and Format B the fewest. This is, in fact, what happened, but the differences were so small that, had Knapp made a statistical analysis, he would almost certainly have found them insignificant. Knapp's results can be seen in Table 35.

TABLE 35

THE RESULTS OF KNAPP'S EXPERIMENT ON PARAGRAPH LENGTH

	COLUMN 1 Number of returns	COLUMN 2 Av. number of paras per return	COLUMN 3 Av. number of words per para
FORMAT A Two columns, elite type, single spaced	65	8.97	123
FORMAT B Narrow margins, elite type, single spaced, massed effect	58	7.43	150
FORMAT C Single spaced, pica type	194	8.72	126
FORMAT D Double spaced, pica type	77	8.88	124
ALL FORMATS	394	8.48*	131*

* With each format equally weighted : that is, as if there were the same number of returns for each format.

(From Knapp (1967 : 35))

The conclusion Knapp draws is that visual aspects do, in fact, play a part in paragraph division. The more strongly marked effect seen in Format B suggests that shape is an important factor in this - that readers do not like wide, short paragraphs, but can tolerate long, narrow ones.

Knapp suggests that an ideal paragraph should not be more than twice as wide as it is long.

Knapp's experiment reveals a tendency, but his failure to apply statistical tests conceals the fact that his results are probably too indeterminate to draw legitimate conclusions. An attempt is made here to apply a different method to the same problem, in the expectation of obtaining more clear-cut results.

3. Experiment 4

While Knapp found that 48% of all subjects divided the passage into seven, eight or nine paragraphs, the spread was very wide, from 4 to 23 paragraphs for Format C alone. He points out that even educated mother-tongue speakers of English may be 'inexpert or confused in the conventions of expository writing', a conclusion supported by pilot experiments on paragraphing conducted by this writer. It was therefore decided to limit the response variations by presenting differently paragraphed passages to subjects for judgment, rather than requiring them to make the divisions themselves. The following hypothesis would be tested :

Hypothesis 5 Experienced readers are influenced by format when making judgments over 'good' or 'bad' paragraphing. They will prefer longer paragraphs when the lines are long and/or single spaced, and shorter paragraphs when the lines are short and/or double spaced.

3.1 Subjects The same 55 subjects were used as for Experiment 1 and 2, divided into three groups of 19, 18 and 18. Again, an attempt was made as far as possible to place an equal number from each department represented into each group.

3.2 Material As only one source passage was to be used, it was important that each subject should be exposed to only one format type. This fact, combined with the small number of subjects available compared with Knapp's experiment, meant that the number of different format types had to be limited to three only. Moreover, it was felt that subjects could not be expected to make accurate choices from more than three possibilities. These three types were as follows :

(i) Wide Format Very narrow margins, single-spaced.

(ii) Normal Format Normal margins, double-spaced.

(iii) Columns Format Two columns, single-spaced, 50% reduction in size.

Although the Wide Format would be the least familiar, none of these types could in any sense be considered abnormal. The Columns Format is almost identical in size and shape to the type of layout found in the daily press, the main difference being the unevenness of the line endings. In any case, it was not considered necessary exactly to duplicate a familiar layout - all that was required was the avoidance of a format so abnormal that it would instantly be rejected.

The Soil Erosion Passage was the source passage used, reproduced in three different versions : Version A

divided into four paragraphs, Version B into eight paragraphs, and Version C into twelve paragraphs. The nine resulting passages can be seen in Appendix 10.

An important consideration was ensuring that each version was equally valid, so that choices made between them would be likely to be based on format type rather than on actual correctness or otherwise. For this, use was made of the tree diagram based on Grimes' Semantic Grammar of Propositions used in Chapter 6 (see Appendix 6).

Version A was paragraphed according to Level 2 of the diagram. The tree was read downwards, and at any movement up to Level 2 or above a new paragraph was begun. This resulted in new paragraphs at Sentences 9 (Level 2), 20 (Level 1), and 27 (Level 2). Sentences 20 and 27 were the author's original choices, although he preferred Sentence 18 to Sentence 9.

Version B was paragraphed according to Level 3, any movement up to Level 3 or above being used as a paragraph division. This resulted in new paragraphs at Sentences 9 (Level 2), 12 (Level 3), 18 (Level 3), 20 (Level 1), 23 (Level 3), 27 (Level 2), and 29 (Level 3). A new paragraph could have begun at Sentence 24, but a decision was made against this in the interest of keeping the final number down to eight. As the movement upwards here is one of only one level, the omission seems justifiable. The last division, at Sentence 29, is the only one not occurring at an upward movement. It occurs, however, at an important node

dominating a large number of propositions, a node chosen, in fact, by 18.95% of the subjects in Experiment 3 according to the weighted analysis.

Version C was paragraphed according to Level 4, any movement up to Level 4 or above being used. This resulted in twelve paragraphs, beginning at Sentence 9 (Level 2), 12 (Level 3), 16 (Level 4), 17 (Level 4), 18 (Level 3), 20 (Level 1), 23 (Level 3), 25 (Level 4), 27 (Level 2), 29 (Level 3), and 32 (Level 4). Sentence 24 seems at first sight a more likely point for a new paragraph than Sentence 25, in that at 24 there is a movement upwards in the tree, even if of only one level. This sentence, however, was considered as a type of bridge sentence, and it was felt that in the interest of avoiding a one sentence paragraph here it should be placed at the end of a paragraph.

3.3 Method Each group of subjects was required to judge between the three different versions of one format type. Group 1 (19 subjects) was given Format A (Wide), Group 2 (18 subjects) Format B (Normal) and Group 3 (18 subjects) Format C (Columns).

The subjects were told that the only difference between the three passages was the paragraphing, and that each method of paragraphing was equally correct. They were then asked to indicate their order of preference between the three versions. The raw results appear in Table 36.

TABLE 36

EXPERIMENT 4 : RAW RESULTS

(a)

Format A (Wide)

CHOICES	VERSIONS 1	2	3
1st	14	5	0
2nd	3	14	2
3rd	2	0	17

(b)

Format B (Normal)

CHOICES	VERSIONS 1	2	3
1st	10	7	1
2nd	2	11	5
3rd	6	0	12

(c)

Format C (Columns)

CHOICES	VERSIONS 1	2	3
1st	5	9	4
2nd	2	6	10
3rd	11	3	4

3.4 Statistical analysis of results In all cases a probability level of 0.05 is required to indicate significance, while a level of 0.01 would indicate high significance.

3.4.1 Chi-square tests applied to each format (See Table 37) Null Hypothesis (H_o) : Taking each format separately, the proportion of first, second and third choices made for each

of the three versions is the same in all cases.

Alternative Hypothesis (H_1) : Taking each format separately, the proportion of first, second and third choices made for each of the three versions differs from case to case.

TABLE 37

CHI SQUARE TESTS APPLIED TO EACH FORMAT

(a)

Wide Format

VERSIONS \ CHOICES	1st	2nd	3rd	Total
A	6.33 / 14	6.33 / 3	6.33 / 2	19
B	6.33 / 5	6.33 / 14	6.33 / 0	19
C	6.33 / 0	6.33 / 2	6.33 / 17	19
TOTAL	19	19	19	57

(b)

Normal Format

VERSIONS \ CHOICES	1st	2nd	3rd	Total
A	6 / 10	6 / 2	6 / 6	18
B	6 / 7	6 / 11	6 / 0	18
C	6 / 1	6 / 5	6 / 12	18
TOTAL	18	18	18	54

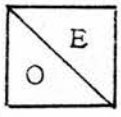
(c)

Columns Format

CHOICES

VERSIONS	1st	2nd	3rd	Total
A	6 / 5	6 / 2	6 / 11	18
B	6 / 9	6 / 6	6 / 3	18
C	6 / 4	6 / 10	6 / 4	18
TOTAL	18	18	18	54

KEY



O = observed figure

E = expected figure

As one might expect from the Knapp experiment, the Wide Format produced the most marked differences, with a X^2 value of 57.17, which is highly significant. For the Normal Format X^2 = 26.02, still highly significant. Both of these formats show a probability of less than 0.001. The Columns Format has a X^2 value of 14.02, but is nevertheless still highly significant ($p = < 0.01$).

The Null Hypothesis can therefore confidently be rejected, showing that there is indeed a marked difference in each format between the subject choices, depending on the number of paragraphs present.

3.4.2 Analysis of residuals (formats) (See Table 38 and Appendix 11) This is a method for identifying the categories responsible for a significant chi-square value, and has already been described (Chapter 8, 3.4.1.4).

TABLE 38

FORMATS : ADJUSTED RESIDUALS

(a) Wide Format

	1st	2nd	3rd
A	**+4.55**	**-1.98**	**-2.57**
B	-0.79	**+4.55**	**-3.76**
C	**-3.76**	**-2.57**	**+6.34**

(b) Normal Format

	1st	2nd	3rd
A	**+2.43**	**-2.43**	0
B	+0.61	**+3.04**	**-3.66**
C	**-3.04**	-0.61	**+3.66**

(c) Columns Format

	1st	2nd	3rd
A	-0.61	**-2.43**	**+3.04**
B	+1.82	0	-1.82
C	-1.22	**+2.43**	-1.22

As one would expect from the chi-square probabilities, the Wide and Normal Formats present the largest number of significant categories. The Wide Format is almost entirely symmetrical, with positive significant categories across

the middle diagonal for the first, second and third positions, and negative categories elsewhere. All categories are significant except B in first position. The most heavily weighted category is that of C in third position, with a positive 6.34, the heaviest weighting in the whole group.

The Normal Format presents the same picture, though with less heavy weighting. The first, second and third positions across the centre diagonal are still significant, but the remaining categories are no longer all negative.

The Columns Format has fewer significant categories, only A in second (negative) and third (positive) positions, and C in second (positive) position. This is, however, a strong reversal from the Normal Format, which is a continuation of the tendency present in the Wide Format.

3.4.3 Chi-square tests applied to each version (Table 39)

Null Hypothesis (H_o) : Taking each version separately, the proportion of first, second and third choices for each of the three formats is the same in all cases.

Alternative Hypothesis (H_1) : Taking each version separately, the proportion of first, second and third choices for each of the three formats differs from case to case.

TABLE 39

CHI-SQUARE TESTS APPLIED TO EACH VERSION

(a)

Version A

(4 paragraphs)

	1st	2nd	3rd	Total
WIDE	10.02 / 14	2.42 / 3	6.56 / 2	19
NORMAL	9.49 / 10	2.29 / 2	6.22 / 6	18
COLUMNS	9.49 / 5	2.29 / 2	6.22 / 11	18
TOTAL	29	7	19	55

(b)

Version B

(8 paragraphs)

	1st	2nd	3rd	Total
WIDE	7.25 / 5	10.71 / 14	1.04 / 0	19
NORMAL	6.87 / 7	10.15 / 11	0.98 / 0	18
COLUMNS	6.87 / 9	10.15 / 6	0.98 / 3	18
TOTAL	21	31	3	55

(c)

Version C

(12 paragraphs)

	1st	2nd	3rd	Total
WIDE	1.73 / O	5.87 / 2	11.40 / 17	19
NORMAL	1.64 / 1	5.56 / 5	10.80 / 12	18
COLUMNS	1.64 / 4	5.56 / 10	10.80 / 4	18
TOTAL	5	17	33	55

KEY

E / O

O = observed figure

E = expected figure

The results of all three tests prove significant, though Version C produced the only highly significant result (X^2 = 18.69, p = < 0.001). The A and B versions show little difference, with X^2 values of 10.80 and 10.32 respectively, and probabilities of less than 0.05.

The Null Hypothesis can therefore be rejected, showing that there is a significant difference in each version between the subject choices, depending on the particular format present. It should be noted, however, that this test is not as reliable as the previous one, in that over 20% of the cells contain expected figures of below 5.

3.4.4 Analysis of residuals (versions) (See Table 40 and Appendix 11)

TABLE 40

VERSIONS : ADJUSTED RESIDUALS

(a) Version A (4 paragraphs)

	1st	2nd	3rd
WIDE	**+2.25**	+0.28	**-2.61**
NORMAL	+0.49	-0.25	-0.25
COLUMNS	**-2.74**	-0.14	**+2.91**

(b) Version B (8 paragraphs)

	1st	2nd	3rd
WIDE	-1.31	+0.08	+1.26
NORMAL	+1.87	+0.50	**-2.41**
COLUMNS	-1.29	-1.24	**+2.55**

(c) Version C (12 paragraphs)

	1st	2nd	3rd
WIDE	-1.71	-0.64	**+2.36**
NORMAL	**-2.39**	-0.35	**+2.76**
COLUMNS	**+3.25**	+0.69	**-3.98**

Significant categories in the A Version are first place (positive) and third place (negative) in the Wide Format, and first place (negative) and third place (positive) in the Columns Format. The table is thus fairly symmetrical.

The B Version has only two significant categories : third position (negative) in the Normal Format and third position (positive) in the Columns Format.

The C Version, which, unlike the other two, is highly significant according to its X^2 value, has far more significant categories, including one of -3.98 (third position, Columns Format), the highest in this group. All categories on the Columns Format, in fact, are significant, the first two positions being positive and the third negative.

3.4.5 Goodman and Kruskal's Lamda Measure of Association

(See Table 41 and Appendix 11) Everett (1977 : 56) points out that no measure of association for a contingency table has yet been produced that is completely satisfactory, as the interpretation of the values obtained is problematic. However, if a group of tables is being considered, it is useful to be able to compare the strength of association between the two variables, even though one figure in isolation is meaningless.

The Goodman and Kruskal Lamda Measure is in actuality a measurement of the predictive ability of one variable in relation to the other, but Everett suggests

that although prediction may not be the purpose of the investigator, predictability does, after all, involve finding a relationship between the variables. Therefore a measurement of prediction is also a measurement of association.

TABLE 41

GOODMAN AND KRUSKAL'S LAMDA MEASURE OF ASSOCIATION

(a) Formats		(b) Versions	
WIDE	0.68	A (4 paras)	0.24
NORMAL	0.42	B (8 paras)	0.17
COLUMNS	0.33	C (12 paras)	0.31

The measurements show that, of the formats, the Wide Format has the closest relationship between the number of paragraphs and the positions awarded by the subjects (λ = 0.68). The versions do not on the whole show such a strong relationship as the formats, but the C Version has the closest association between the format and the position awarded (λ = 0.33). (0 = no predictive association, unity = complete predictive association).

3.5 Interpretation of results (See Figures 29 and 30)
Looking first at the formats separately, as Knapp did, it is evident that the Wide Format produces the most definite results. Here, the subjects very clearly prefer the four paragraph version, (A), with eight paragraphs (B) second

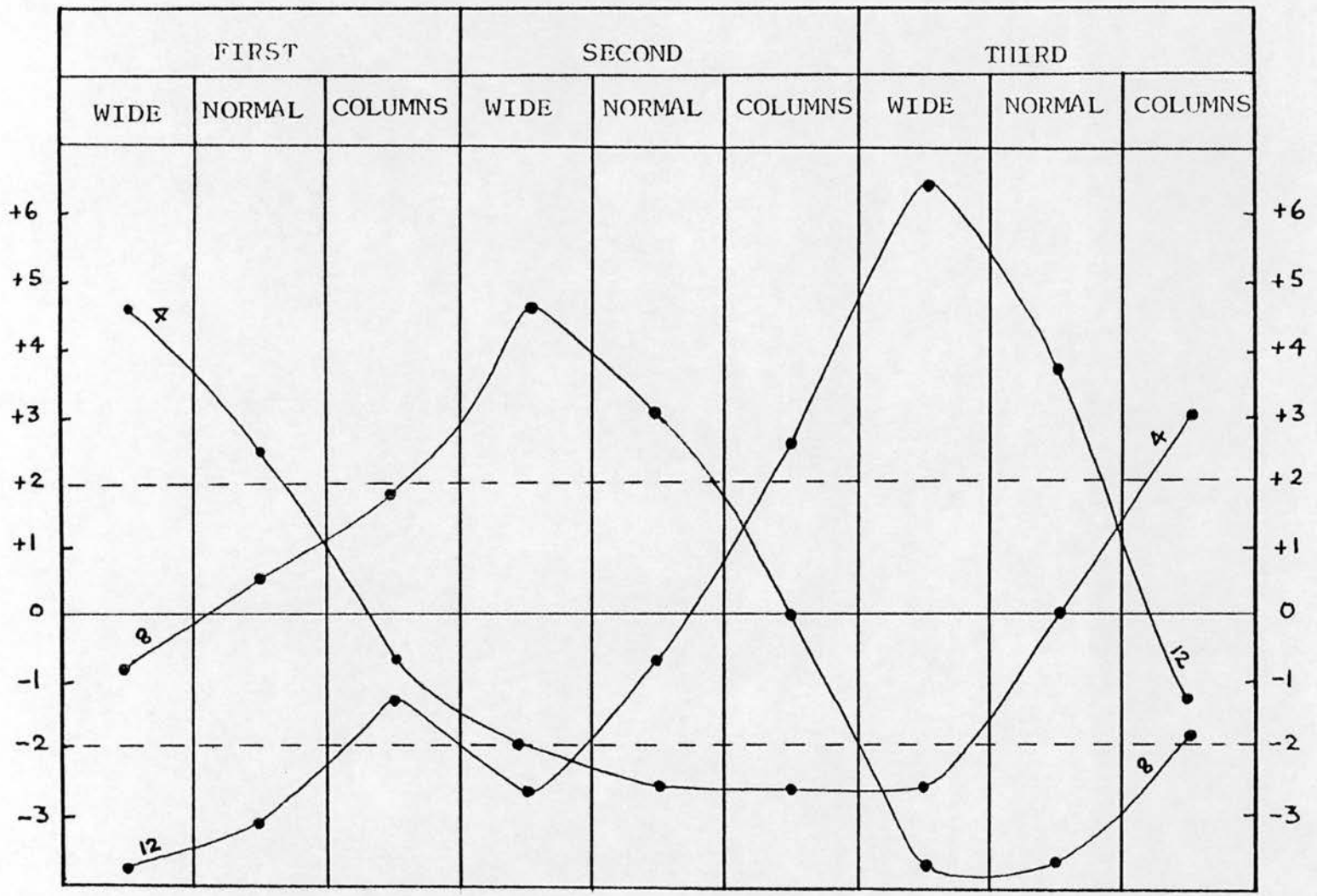


FIGURE 29 ADJUSTED RESIDUALS : FORMATS PLOTTED AGAINST VERSIONS

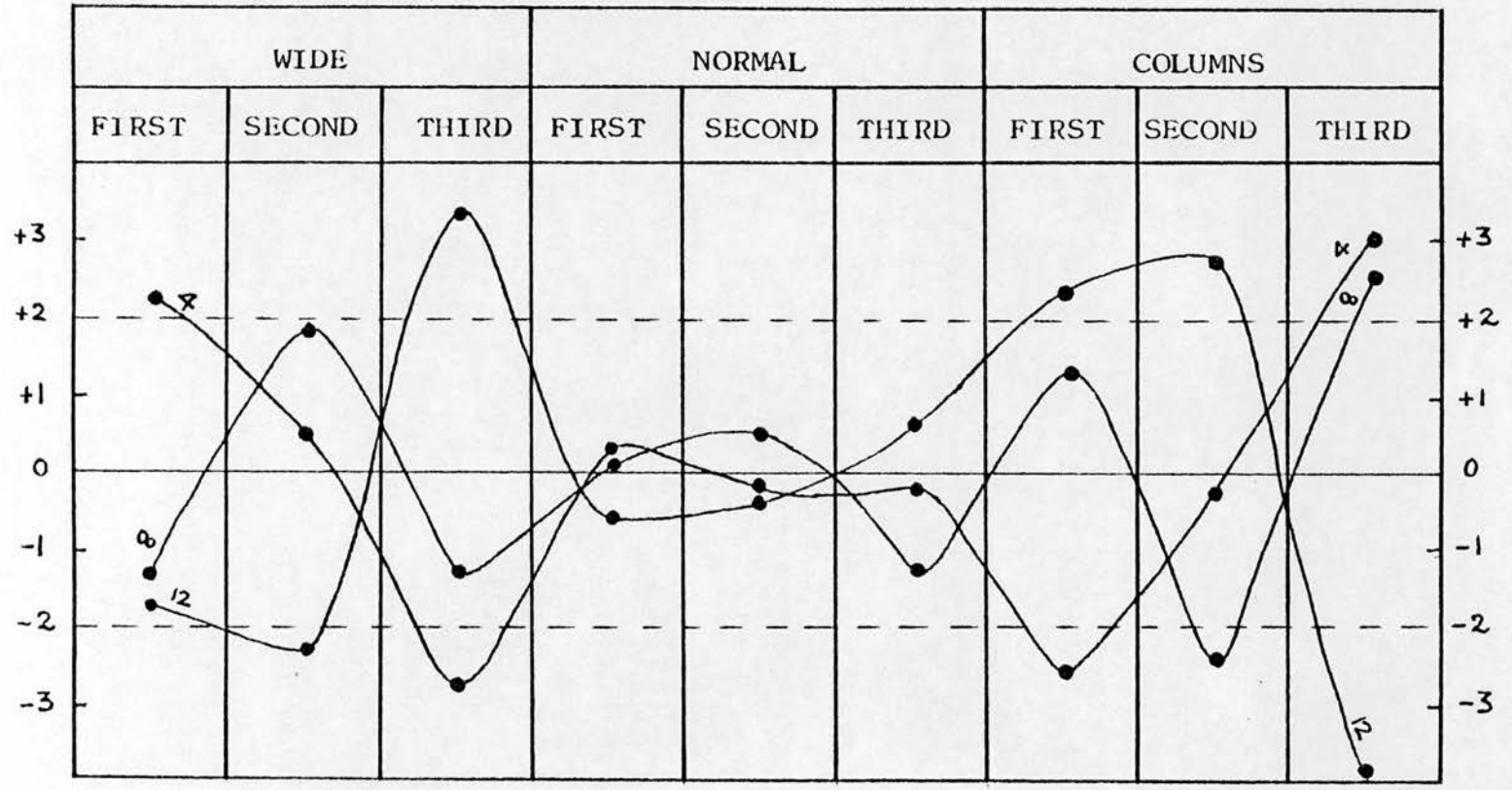


FIGURE 30 ADJUSTED RESIDUALS : VERSIONS PLOTTED AGAINST FORMATS

and twelve (C) third. The position of the twelve paragraph version in third place is, in fact, the most heavily weighted choice made on any version, with an adjusted residual value of +6.34. This corresponds favourably with Knapp's results.

The Normal Format does not differ markedly, the same versions holding the same positions, but in every case with the totals reduced. This can be clearly seen from the pattern of adjusted residuals, where the values for first, second and third position have been roughly halved, and the remaining categories are no longer all negative. It is clear that while subjects still prefer the four paragraph version (A) their preference is not so marked, and the gaps between the versions is narrowing. The lengthening of the paragraphs in proportion to their widths seems to have had the result of producing a greater variety of response from the subjects.

The Columns Format continues the trend to the point where there is a significant reversal. The tide has now turned against the four paragraph version to the extent that it is no longer in first position - it now stands as the main choice for third position, with an adjusted residual value of +3.04. The twelve paragraph version (C) does not take first place, however, but moves into second position, leaving first position for the eight paragraph version (B).

Figures 29 and 30 deal with all the data combined.

In Figure 29 the adjusted residuals of the formats are plotted according to the positions of the versions. Here the versions follow each other in succession in a wave-like pattern. In the Wide Format, the four paragraph version tends to hold first position, the eight paragraph version second position and the twelve paragraph version third position. Although the trend is clear, it is the Wide Version here which presents the most extremes, having, as it does, the greatest number of significant residuals.

To emphasize visually the positive effect of the Columns Version, it is necessary to turn to Figure 30, where the adjusted residuals of the versions are plotted according to their positions within the formats.

Two important points stand out here. First, we have the increase in the number of first and second choices for the twelve paragraph version (C) in the Columns Format. The graph does not mean, of course, that this version supplied more first choices than the other versions - in reality it ranked third numerically, with a total of four, as opposed to nine for Version B and five for Version A. What it does mean, however, is that of a total of five first choices spread over all three versions, no fewer than four were from the Columns Version, a statistically significant figure.

The second important point is that it is immediately obvious from this graph that the Normal Format holds an intermediate position. There is very little difference

here between the figures for the three versions, not one of the adjusted residuals being significant. On either side of the Normal Format the Wide and Columns Formats present almost exact mirror-images of each other, the only major difference being the twelve paragraph version's predominance in second position rather than first in the Columns Format.

Considering both the statistical tests and the above interpretation, the Null Hypothesis can therefore be rejected. As far as this experiment is concerned, the readers have been strongly influenced by the formats in making their judgments over preferred paragraphing. They tended to select longer paragraphs when the lines were long and single spaced, and shorter paragraphs when short and single spaced. When the lines were medium to long and double spaced their choices were less definite, but tended to lie at some point between the two above extremes. This result suggests that the paragraph cannot be considered as an independent organic entity, but must be viewed more as an orthographic device dependent to some extent on the idiosyncratic preferences of the writer.

CHAPTER 9

A THEORETICAL SUMMARY

1. Introduction

The basic assumption of this work has been that traditional rhetoric provides a valuable source of information and ideas concerning paragraphing and text structure in general, but that these ideas have never been satisfactorily systematised and unified. An attempt has been made here to verify some of the rhetorician's prescriptive assertions experimentally, and to place them within a hierarchical model.

An outline of the model for expository prose structure and paragraphing as envisaged by this writer has already been presented (see Chapter 4, Section 2), so will not be stated yet again. Some general points, however, should be made about the experiments, but first a few remarks are needed concerning some of the problems encountered in these experiments.

2. Some difficulties resulting from the use of text

It cannot be too strongly emphasized that many problems arise with the use of text in experimentation. It could, indeed, be argued that written discourse, or at least authentic written discourse, is in no way suitable raw material for the experimental approach,

as it can only be viewed from an individual and subjective standpoint, the knowledge and attitudes the reader brings to the text being largely unpredictable.

When artificial texts are especially constructed, the main difficulty is the control of a wide variety of variables. For example, in Experiment 1 differential subordination and embedding between the passages may have influenced the subjects when they made their choices, and in Experiment 2 even quite a slight difference in the length of the short sentences between the two passages could have affected the subjects' judgments. Any variation in the pattern of information structure between the passages could also have adversely affected the results. The identification and control of such extraneous variables is a far greater problem when dealing with text than it is in similar sentence-based experiments.

The use of Grimes Semantic Grammar of Propositions, or indeed of any other similar model, as the basis for the construction of textual tree diagrams also presents difficulties, this time difficulties of interpretation. The indication of rhetorical function is often, indeed with some writers usually, implicit, with the result that at the higher levels of the tree various interpretations become possible. Indeed, it could be said that only

the author himself can produce a representation of the content structure with any real reliability.

This, however, is a problem that the discourse analyst must learn to live with. It is part of the nature of communication that interpretation involves guesswork on the part of the listener/reader (See Goodman 1967), and therefore any kind of analysis will involve a certain amount of subjectivity. Expository prose should be the one area where relationships are made the most explicit, where one of the main aims of the writer is painless comprehension on the part of the reader. Thus, although the use of this method of hierarchical structuring is not ideal, it should be acceptable when applied to a well-organised and sign-posted exposition.

If such an analysis proves particularly difficult, then this will reflect the problems the reader faces in comprehending the text. It is significant that of the three passages treated in this way the one that lent itself most readily to such a representation (Passage A) presented the least difficulty to the subjects required to insert paragraph boundaries. Passage C, on the other hand, was more difficult on both counts. This suggests that a very poorly structured student essay might prove quite impossible to transform into a tree.

3. The experimental work

Some general points will now be made as to the theoretical conclusions to be drawn from each of the four experiments. A few points will also be made concerning the topic sentence and the bridge sentence.

3.1 Experiments 1 and 2 were an attempt to find some evidence to support the prescriptions of traditional rhetoric concerning sentence length. Much more attention could have been paid to the sentence as an information block in its own right, as, for example, the important question of the amount of coordination, subordination and embedding present. This, however, would have taken us into consideration of the internal structure of the sentence, and thus would have trespassed beyond the assigned boundaries of this work.

Experiment 1 indicates that the experienced reader reacts against what he considers to be a succession of excessively long or excessively short sentences. He prefers passages which consist of sentences of medium and varied length, but no clear preference is shown between these. This indeterminacy, however, may be the result of the shortness of the experimental passages themselves. It is possible that a succession of medium-length

sentences over a much longer stretch of text would have been less favourably received. On the other hand, the use of intermittent short sentences is a very common idiolectal feature rather than a general characteristic of all writing, so certain subjects might always prefer medium-length sentences to those of varied length.

Experiment 2 shows that where sentence-length is varied the short sentences tend to occur at changes of topic, ie at transitions between medium or higher level information blocks. This may simply be that the short sentence signals something of special importance, and among the most important assertions in the text will be those that are macro-structural in function (van Dijk, 1977). A short sentence used in an inappropriate position may misguide the reader by highlighting a proposition which plays only a subsidiary facilitating role. The higher level topic sentences, ie the topic sentences of paragraphs and groups of paragraphs, would tend to be longer, as they dominate a larger section of the hierarchy, while more obviously short sentences would often be the topic sentences of lower-level blocks (sub-topic sentences in traditional rhetoric). Thus the short sentence may be used as a signal to indicate change of topic.

3.2 Experiment 3 Here the paragraph is related to the hierarchical structure of the text. Grimes' Grammar of Semantic Propositions is used to represent the texts as trees, and it is shown that paragraph divisions tend to be made where there is a movement upwards in the hierarchical structure. If it can be assumed that the highest levels of the tree are the most general, then the findings of the experiment can be related to Christensen's 'levels of generality'. Christensen claims that a new paragraph can begin where a succeeding sentence is neither coordinate nor subordinate with that preceding it, and thus the movement could be up the tree or across it. According to the results of this experiment, if movement is across the tree the probability of this movement will depend on the number of propositions the node dominates.

Again, as in Experiments 1 and 2, it is only possible to speak in terms of tendencies and generalities. No rules can be formulated as to exactly where a paragraph boundary will be placed - it is only being claimed here that a paragraph is more likely to begin in one place than in another if the vertical distance of the movement is greater and the node higher, or the number of propositions dominated by the node greater. The writer has a certain amount of choice as to the positioning of his paragraph boundaries, but this choice is constrained

by the overall structure of the text and by the density of propositions beneath particular nodes. For example, he may, if he wishes, and is indeed likely to, begin a new paragraph at a point where there is a movement upwards in the hierarchical structure, but he is very unlikely to do so between two lower-level propositions supporting a general statement in the node above.

As far as this part of the experiment is concerned, nothing really new is being claimed. Nevertheless it reflects what is intuitively real to the mother-tongue speaker, namely that the first sentence of a paragraph is normally a general statement, which is followed by supporting statements of various kinds, and that the general statements are themselves related at higher levels. The experiment attempts to some extent to formalise this intuitive reality. Whereas Christensen limits his attention to single paragraphs in isolation, the aim here is to show how the paragraph organisation relates to, and depends on, the text in its entirety.

Experiment 3 was also concerned with the type of signal used to mark movement to a new information block. Unlike Becker, who divided his signals according to the three hierarchies of lexicon, grammar and phonology, this writer has chosen to differentiate between intrinsic markers of content and extrinsic

markers of cohesion. This classification is related to ease of manipulation by the writer, in that the extrinsic signals are more independent of the semantic content, and can be consciously inserted into or omitted from the text.

The experiment indicates that the readers tend to more easily recognise a transition if it is marked by an extrinsic signal. The most significant is the text reference signal, which seems to occur most frequently at intermediate and higher levels in the hierarchy. The highest level text references would be considered by van Dijk (1977) as manifestations of the pragmatic macro-structure of the text, an example being the macro-speech act "We may sum up by saying that ...." (Passage A, S20). Contrast words may appear anywhere, but in these passages tend to function at a fairly low level. Short sentences have already been mentioned as characteristically functioning as sub-paragraph topic sentences.

Thus Experiment 3 suggests that more signals, and in particular more extrinsic signals, tend to occur at the higher levels. The type of extrinsic signal used varies somewhat from writer to writer, but generally text reference signals are the most effective, with the use of the short sentence becoming more marked at intermediate levels within the paragraph.

3.3 Experiment 4 had the aim of showing the flexibility of the paragraph boundary. Traditionally a paragraph was considered as something real in itself, an organic whole, whereas the suggestion here is that it is no more than a rather powerful form of explicit signalling, based not only on overall text structure but on a number of external factors, such as length of line, type of print, and size of page as well as the preference of the individual writer. For example, if the lines are long and/or single-spaced paragraphs will tend to be longer, and if the lines are short and/or double-spaced they will tend to be shorter, with the proviso that the writer is free to vary this pattern.

The aesthetic appearance of the paragraph on the page is undoubtedly an important factor here. Knapp (1967) has suggested that a paragraph should not normally have a width which is more than twice its length. This is supported by Borissavlietch (1958) (quoted in Watson 1973 : 107), who described an experiment in which subjects were asked to state their preferences between a number of shapes. The majority chose a rectangle with a width a little more than one and a half times its length, a shape very close to the long-recognised 'Golden Mean' of the artist.

Of course, it must not be forgotten that many different variables, with reference to both the text and the reader, are involved. If the material could be considered difficult for the target readership, then the short paragraph might quite well be used, as can be seen by comparing paragraph lengths in mass circulation newspapers with those in the quality press. Textbooks written for primary-school children will tend to have shorter paragraphs than those intended for university students, the variables here being the age and level of the projected readership and the density of information being presented. Nevertheless, the experiment does indicate the importance of the paragraph boundary as an extrinsic signal which can be manipulated by the writer to indicate the hierarchical organisation of the text, and as a feature which is especially important by virtue of its eye-catching quality.

4. The bridge sentence and the topic sentence

These two features were not experimentally investigated, but an attempt was nevertheless made to present some ideas based on the experimental passages and the subjects' responses to them. These tentative conclusions might hopefully form the bases of future hypotheses.

4.1 The bridge sentence. Even in traditional rhetoric the bridge sentence was seldom considered. Here it has been defined as a sentence which could equally well be categorised as part of the information block preceding it, or as part of that succeeding it. Bridge sentences were identified in the passages in places where the subjects were uncertain in their choices between two consecutive sentences.

The bridge sentence seems to follow a commonly recurring pattern. Its first part usually refers back to the immediately preceding information block, and its second part states the main topic of the succeeding block. Often a word or phrase expressing a relationship of contrast or result is to be found in the middle of the sentence.

The distribution of given and new information seems to be an important factor in their identification. Medium to higher level information blocks tend to begin where 'new' and 'theme' coincide, but the bridge sentence follows the unmarked pattern of 'theme' corresponding to 'given', and is succeeded by a sentence of the same pattern.

4.2 The topic sentence. Traditional rhetoricians and teachers of composition have taken the sentence as their basic unit. They have viewed the topic sentence as the sentence expressing the main theme of

the paragraph, and normally placed in first position. Thus the topic sentence as traditionally conceived was a necessary component of the paragraph, and occurred at middle and higher levels in the hierarchy.

Christensen attempted to incorporate the topic sentence into his model when he defined it as the top sentence of a sequence of structurally related sentences. Thus it could be seen as a general statement supported by succeeding statements which were either coordinate or subordinate to each other, the whole group of sentences together forming a paragraph.

One problem immediately evident is that the topic may often be expressed as one part only of a sentence, the rest consisting of supporting propositions manifested as subordinate clauses. Thus it is more reasonable to think in terms of topic <u>propositions</u> rather than topic <u>sentences</u>.

The suggested definition of a topic sentence, then, is that it is any sentence that contains no more than three propositions, and whose main proposition is a topic proposition, a topic proposition being any proposition which directly dominates at least three other propositions. This ensures that all topic sentences thus defined are fairly short, in some cases short enough to act as signals of change of topic.

A topic sentence, however, need not only represent the topic of an orthographic paragraph. Similar sentences at a lower level within the paragraph occur (eg Passage A S8, Passage B S13, Passage C S28), and at a higher level a sentence may express the topic of a group of paragraphs (eg Passage A S20, Passage B S2, Passage C S3). Moving below and above the sentence, a topic may be expressed by the main clause of a sentence, or by a paragraph, a group of paragraphs or a chapter.

According to van Dijk (1977) topic propositions, together with such features as titles and sub-titles may be viewed as macro-propositions, as representations of the macro-structure of the text. A macro-proposition is "a proposition entailed by the sequence of propositions underlying the discourse (or part of it)" (van Dijk 1977b : 137), and it is obtained from the micro-structure by the application of mapping rules. Topic sentences could thus be considered as macro-propositions, which would aid the reader in his interpretation of the text by making it unnecessary for him to apply these mapping rules.

5. An analysis of two representative examples of student writing

As an example of the way in which such a view of the overall structure of expository prose

might be applied, the two essays already presented in Chapter 1 (Figs 1 and 2) as typical examples of student writing will be briefly considered. This analysis should be read in relation to the conclusions presented in section 2 of Chapter 4 and sections 3 and 4 of this chapter.

Space does not allow detailed treatment, but each essay will be considered generally as regards higher level organisation and paragraphing. In order to facilitate the analysis an attempt has been made to produce a tree diagram of the content structure of each essay (Figures 31 and 32). These diagrams may be viewed with an element of scepticism, as in some places a complete omission of bridging elements and signals makes their interpretation so difficult that the connections suggested are decidedly doubtful. Nevertheless, the essays were written to be read, even if only by the supposedly all-knowing teacher, and these are possible interpretations. The difficulties encountered in transforming the texts into trees are therefore an indication of the organisational inadequacies of the writing, and reflect the problems of interpretation faced by the reader.

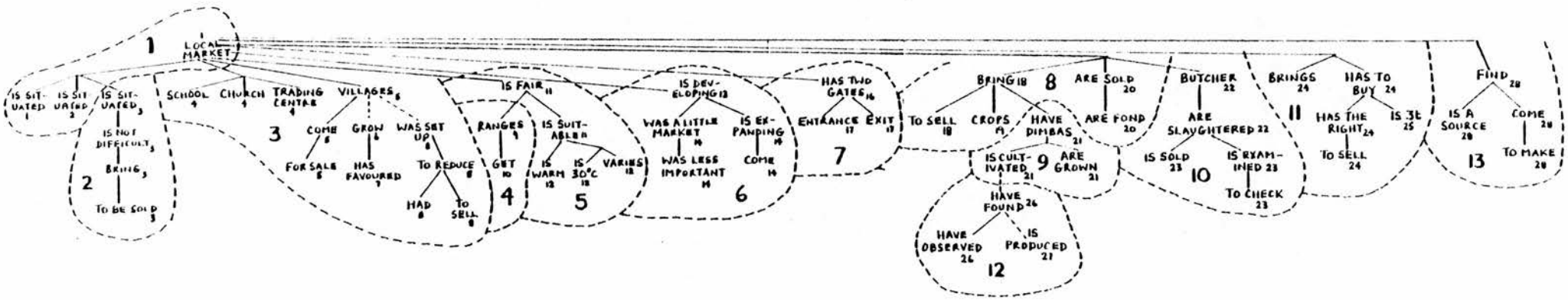


FIGURE 31 TREE DIAGRAM OF AN EXAMPLE OF STUDENT WRITING (SEE FIGURE 1)

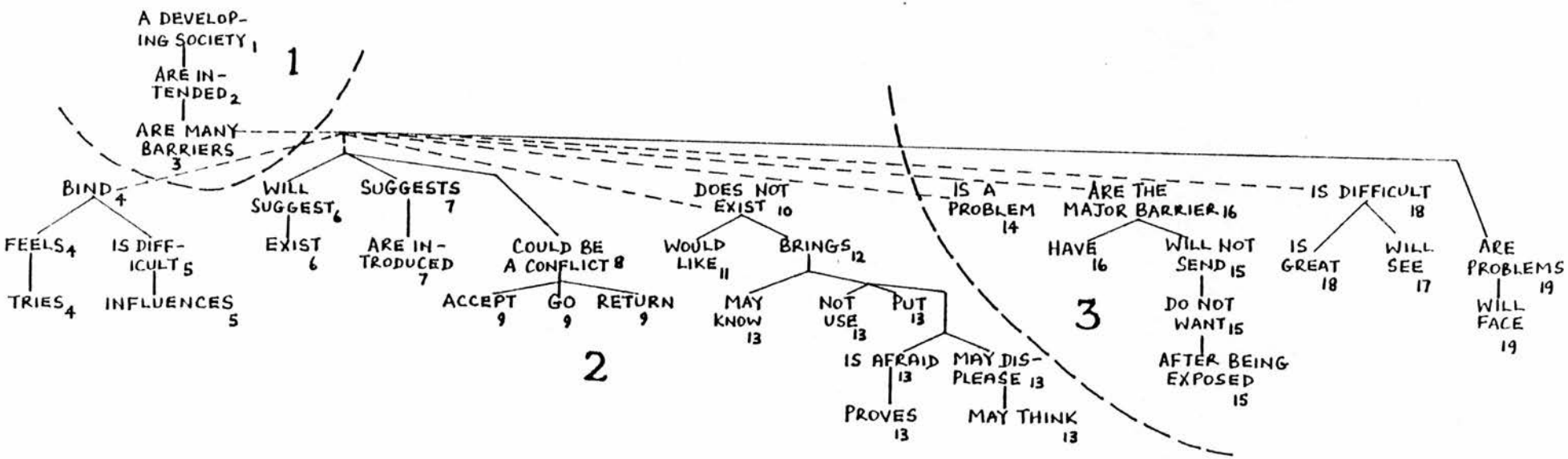


FIGURE 32 TREE DIAGRAM OF AN EXAMPLE OF STUDENT WRITING (SEE FIGURE 2)

5.1 Passage 1 (Figures 1 and 31) The hierarchical structuring of the essay is rudimentary, consisting of a chain of mostly poorly developed information blocks at the same level, with no differentiation made between more and less important areas. For example, the seemingly irrelevant point that the market has two gates is given the same prominence and appears at the same level as more significant statements concerning the climate or the goods on sale.

Most of the paragraphs are treated as isolated units, the only extrinsic link used being the connector 'also' in Paragraph 10. There is a general shortage of discourse adjuncts and no bridge sentences or bridge paragraphs.

Each paragraph tends to be short, with an average length of only 38.23 words. Sometimes this is because separate information blocks have not been sufficiently developed (Paragraphs 7 and 11), but more often a higher level information block has been broken up into needlessly small units (Paragraphs 1 and 2, 4 and 5). An improvement cannot always be made simply by combining paragraphs. In cases where information blocks are under-developed, either more content should have been added, or, if this gives undue prominence to less important topics,

some method should have been found of re-grouping the blocks under a common super-ordinate topic.

It is difficult to refer to topic sentences in a text where the higher level information blocks are so small and undeveloped, but at least four can be identified. Sentence 11 is the topic sentence for Paragraphs 4 and 5, but has been wrongly positioned. Sentences 13, 18 and 22 can be considered as topic sentences, the first also being a short sentence. With such under-developed information blocks it is inevitable that the topic sentences all occur at the same level, and relatively short sentences are few and far between.

The paragraphs are strung together in a haphazard manner, becoming more random as the text proceeds, so that the reader is given few clues as to what will follow from one paragraph to the next. Thus the high number of very small disconnected paragraphs results in an over-large element of surprise, if not mystification, for the reader, especially towards the end of the essay. Higher level coherence breaks down completely when the reader is not clear what the link is between adjacent information blocks. This mainly results from the writer's appearing to place equal importance

on all points made, which is in turn the result of a weak or non-existent system of conventional ordering, and the use of very few levels. For example, it is difficult to accept the leap from Paragraphs 6 to 7 and then to 8, where a seemingly low level and rather specific topic (7) is placed between two more developed topics beginning at a more general level (6 and 8). These have a clear relation to each other but not to the intruding paragraph.

5.2 Passage 2 (Figures 2 and 32) From the title the reader would expect a classification of the various barriers identified by the writer, with a brief explanation of each. He might expect them to be placed in some kind of logical order, and to be followed by a brief summary. In actuality, rather more guesswork than is reasonable to expect is necessary in order to identify the barriers listed, as they are linked together in random order with an almost complete absence of extrinsic signalling. Signalling is all the more essential here, as numbered headings and sub-headings to indicate topics and sub-topics are not traditionally acceptable in an essay. Thus a text with few grammatical errors becomes in places almost incomprehensible.

The passage has been divided into only three paragraphs, with an average length of 129 words. This figure conceals the fact that the shortest (initial) paragraph has only 32 words, and the longest as many as 143. Thus the passage could have been further sub-divided. The apparent reason for the division between the second and third paragraphs is the division of the six barriers into two groups of three barriers each. However, as the arrangement of the barriers is random, the real reason is more likely to have been the writer's feeling that the second paragraph had become over-long.

The third paragraph contains information blocks that should on no account have been placed together. The last sentence within the paragraph represents a general conclusion to the entire list of barriers, not simply to the three barriers mentioned in the same paragraph. It should therefore have been placed in a paragraph on its own, in which case it would have needed some further development.

As in Passage 1, the information blocks are almost completely unsignalled and the paragraphs appear as units isolated from each other. There are no bridge sentences. The last sentence of Paragraph 1 leads the reader to expect the next

topic to be a list of barriers, but the first sentence in Paragraph 2 in no way makes it clear that this assumption is correct. In fact, the word 'barrier', or even the word 'problem', is not used at all in Paragraph 2, so that the reader is faced with an almost clueless puzzle, which requires considerable ingenuity to even partially decifer. Similarly, Paragraphs 2 and 3 are completely unconnected, and present a particular problem in that there seems no valid reason for a bread to occur at this point.

In the second and third paragraphs a few possible topic sentences occur, but usually the incoherence of the text makes their recognition as such not always immediately evident, thus defeating their purpose as signals. The most fully developed and comprehensible section, from Sentences 10 to 13, contains two topic sentences (10 and 12), which are also short sentences and occur at different levels. The first sentence of Paragraph 1 is also a short sentence, and the topic sentence of a quite effective introduction. It fails in its function as the overall topic sentence however, because of a lack of a clear and adequate connection between the introduction and the two succeeding paragraphs.

The theme sentence of a paragraph, even if it cannot be considered as a topic sentence, encourages the reader to have certain assumptions concerning what will follow. The theme sentence of Paragraph 3 promises information about attempts to persuade the farmer to use fertilizer, and this promise is especially forceful by the fact that it is expressed in a short sentence (that is, short in relation to an exceptionally long preceding sentence of 63 words). Instead the writer turns immediately to a new topic, the refusal of parents to send their children to school. Coherence breaks down between the sub-topics which represent the different barriers, in some cases because of an absence of topic sentence and in others an absence of text reference signals. In the example just given, it is difficult to see how the gap between the two sentences can be bridged. Presumably these sentences refer at a higher level to different barriers, but even the nature of the first barrier is obscure. Is it the conservative attitude of 'traditional man', or is it the fact of his not believing in fertilizer? One cause of difficulty here may be the writer's inter-changeable use of the lexical items 'barrier' and 'problem'.

As in Passage 1 there is an almost complete absence of discourse adjuncts. Some additive

adjuncts, such as also, too or in addition to, would have been useful to indicate the successive barriers.

5.3 The approach of these writers to text chunking

The above analyses perhaps go some way towards explaining how student writers who can function reasonably well at sentence level nevertheless fail to produce coherent and comprehensible text.

Both writers are aware that paragraphing is necessary, but follow different strategies in this respect. The writer of Passage 1 takes the view that whenever there is a change of topic, a new paragraph should begin. As the topics follow each other in a disordered chain, his paragraphing does the same, each paragraph consisting of an individual information block or part of a block of one to three sentences in length and with very little internal development. The writer of Passage 2 has probably been taught that every essay should contain at least three paragraphs, including an introduction. He therefore writes an introductory paragraph, and then throws the different items of the subsequent classification together into two further paragraphs. Thus he has chunked several information blocks into each paragraph. Unfortunately a lack of signalling

and inadequate information structuring renders the transitions between these blocks almost impossible to decipher. Also his paragraph boundaries are incorrectly related to the hierarchical structure of the text, with the result that they confuse rather than enlighten the reader.

An important reason for the difficulties both passages pose is that neither writer explicitly signals transitions between information blocks, either between or within paragraphs. The information blocks are inadequately developed, so that the texts as a whole lack depth, and are, in Christensen's words, 'thinly textured'.

It is obvious that if these are average student essays at these levels all is not well with the teaching of writing. Some suggestions for the pedagogical application of the findings of this work will be made in the next and final chapter.

CHAPTER 10

SOME PEDAGOGICAL APPLICATIONS

1. Introduction

The aim of this work is not primarily to present new pedagogical ideas. Instead an attempt has been made to justify a familiar but unfashionable approach by demonstrating that some of its prescriptions are factually based and can be formalised. The subject of special concern has been paragraphing and the chunking of text, this seeming to be a particular problem area for second language writers. In doing this, however, certain insights have been obtained concerning specific pedagogical problems, and these will now be briefly considered.

2. Sentence length : Experiments 1 and 2

It is common practice to advise students to avoid monotony of sentence length. Advice, however, is seldom given as to how to achieve this - vague talk about 'rhythm' and 'balance' may be intuitively comprehended by a mature mother-tongue speaker, but means little to most second-language learners with scant experience of texts.

On the whole Malawian students at early tertiary level commonly use sentences of medium length interspersed with occasional short sentences, though

some of the weaker ones tend towards long and often loosely coordinated sentences such as Sentence 13 in Figure 2. This generalisation relates to Hunt's conclusion that as children mature they tend to write longer clauses with increasing subordination, until they reach a point where their organisation of information becomes more "intricate" and such clauses are reduced to phrases (Hunt, 1970). Thus, as a learner's English progresses, he may also pass from the stage of simple kernel sentences through coordination and subordination to reduction, and hence move from over-short, to over-long to medium length sentences. This is merely a suggestion, but is a field which would repay further investigation.

Thus it seems that there may be a natural movement towards sentences of medium length. The problem therefore lies not so much with sentence length in isolation as with the placement of short and long sentences in relation to each other. Many good mother-tongue speakers use the short sentence at change of topic quite intuitively. Indeed, most of us are so unaware of this function that in no traditional rhetoric or college composition handbook examined for this study has it been found necessary to teach methods for the most effective placement of short sentences. Nevertheless this is one of the

reasons for the complaint of many teachers of second language learners that their writing does not read fluently, that the ideas being expressed do not 'flow' as they should.

One reason for the scarcity of short sentences in such writings is the common failure to develop ideas adequately. If information blocks are poorly developed, as in Passage 1 (Figure 1), each consisting of two or three sentences with minimal subordination, then there is little possibility of much contrast in sentence length. Sub-topic sentences in the traditional sense will not be very evident if the tree structure of the passage is shallow.

For pedagogical purposes, a short sentence could be considered as a sentence which consists of only one main clause, or a 'kernel sentence' as defined by Tufte (1971). The short sentence is a <u>possible</u> rather than a <u>necessary</u> signal of topic change, so its function in this way should be made evident to the student rather than actually prescribed to him. Exercises involving the reconstruction of short texts from a sequence of 'kernel sentences' of the type used by Dautermann (1969) and Hunt (1971), and in Experiments 1 and 2 of this work (see Appendix 2c) could be utilised. Such exercises are common for teaching the use of different kinds of subordination and coordination within the sentence, but could be

extended for practice in producing sentences of appropriate lengths.

3. The paragraph in relation to the hierarchical structure of the text : Experiments 3 and 4

The teaching of writing has usually proceeded in a 'bottom-up' direction, ie beginning with the sentence and progressing via the paragraph to the complete text. This has resulted in the conception of the paragraph as an isolated unit, as, indeed, the sentence has also been traditionally viewed. School textbooks have been devoted entirely to paragraph writing, with no suggestion as to how such artificial paragraphs may be developed into longer pieces of discourse. (eg Jupp and Milne (1972), Chaplen (1970)).

If the main function of the paragraph is to indicate hierarchical structure within the text, then it cannot exist in isolation from the text. A student who writes a 'paragraph' as an exercise in the English class is in reality producing not a paragraph but a mini-text, which itself contains information blocks but is too short for these to be indicated by further sub-divisions. The danger of such exercises is that the student learns to view the paragraph as a unit of language complete in itself, and when he progresses towards full-length

texts omits to provide signalling devices or bridges between their constituent paragraphs.

This can be avoided, or at least ameliorated, by as soon as possible, and certainly by the intermediate level, adopting a parallel 'top-down' approach, in which the student is presented with authentic expository text for analysis and discussion. Such texts should preferably be extracts from subject textbooks the student is currently engaged in studying, or with a mixed group they might be selected easy articles from appropriate journals (eg. 'The New Internationalist'). Whichever are used, exact copies should be available, to indicate how the format influences paragraphing and can often be an aid to comprehension (by the use of different type-faces, headings and sub-headings, illustrations and diagrams, etc.)

If a text of the type under consideration usually consists of layers of general statements at different levels, each generality associated with various kinds of supporting statement, then it is necessary for the student to be able to differentiate between general and specific. Such an understanding is obviously vital for the construction of adequate topic sentences. However, the ability to identify generalisations cannot be taken for granted when the student belongs to a non-Western and non-literate

cultural group, as is the case in many parts of sub-Saharan Africa. Exercises are therefore needed for this, as also they are needed for training students in the ability to recognise general statements at different levels of generality.

Such exercises are rare in published materials, but can be found in a few recent textbooks. For example, in "Reading and Thinking in English" Book 3 an exercise appears in which three groups of sentences are presented, a supporting example from Group 1 being a generalisation in Group 2, and a supporting example in Group 2 being a generalisation in Group 3. Students are required to identify the generalisation in each group, and to notice how the different levels of generality depend on the context. Far more attention needs to be paid to such areas - it is too often assumed that a student's difficulty is linguistic when it may in fact be largely conceptual.

A greater awareness of the hierarchical organisation of this type of text can be encouraged by the use of flow chart analyses, the construction of flow charts being useful both for the analysis of published text used as models and for the students' own writing. Again, however, care needs to be taken that students from non-literate cultures are

sufficiently familiar with the techniques of diagrammatic representation, before assuming that these can automatically be used as learning aids.

The 'top-down' approach can equally be applied to production. One possible method is that of expansion from the summary. It was claimed in Chapter 7 that successive topic sentences would form an adequate summary of a text; this suggests that if students are given such a summary they should be able to expand it into a full text again by adding appropriate supporting detail.

Such an approach can also be used to illustrate the close structural relationship between the paragraph and the essay pointed out by some traditional rhetoricians. (eg Hulbert and Hulbert, 1929). If a paragraph is to be seen as a collection of lower level information blocks, then each information block can be expanded into a full-length paragraph in its own right by adding further supporting detail. Such an approach is taken in Imhoof and Hudson's recent textbook for tertiary level students, "From Paragraph to Essay," where each of six sentences of one particular paragraph is expanded, to form an essay six paragraphs long (Imhoof and Hudson, 1975 : 12-15).

The hierarchical structure of an expository essay is often indicated on the preliminary plan

by means of topic headings, but it is usually better to require the student to construct topic sentences which can later be developed by means of supporting statements. If the plan is genuinely hierarchical and not merely a list, this method might usefully indicate to the student that a topic sentence need not necessarily be the main sentence of a paragraph - it could as easily be the topic sentence of a group of paragraphs or of an information block within the paragraph. Discussion should take place before writing as to where paragraph boundaries are to be placed.

This brings us to the question of paragraph length. Teachers commonly advise students to develop each heading into a separate paragraph, a procedure which often results in widely varying paragraph lengths. If discussion also takes place as to which headings may be grouped together according to the planned hierarchical structure of the essay, then very short paragraphs for lower level topics can be avoided.

Failure to develop topic statements adequately often results in paragraphs that are too short, this being compounded by a lack of overall hierarchical organisation. Such a lack of development may reflect a shortage of information,

on the part of the student, which may be rectified by more preliminary preparation in class. Provided the student writer knows enough about his subject, there should be no problem over paragraphs that are too short.

For the average student, however, paragraph length is not a significant area of difficulty. Watson (1973) suggests that response to proportion may well be independent of culture, in that it is probably based on the common human distance between the eyes. Thus it is possible that the aesthetic appeal of particular paragraph shapes will ensure that most writers will normally produce paragraphs of acceptable length.

4. Extrinsic linking devices : Experiment 3

Although it is possible that many good writers, even of expository prose, use few extrinsic signals apart from that of paragraphing itself, we cannot expect the inexperienced second language learner to emulate their example. It is inevitable that most will write fairly incoherently, mainly because of poor control of information structuring procedures, but an ability to use appropriate signals, including short sentences, will go a long way towards enabling them to attain an acceptable level of comprehensibility.

It is now common, even in Malawian secondary schools, explicitly to teach the correct use of logical connectors, but too often the focus is on connections between adjacent sentences to the exclusion of connections between higher level chunks or groups of sentences. This omission was indeed one of the major weaknesses of Christensen's work, in that he only considered the relationship between a sentence and that which immediately preceded it. Thus exercises which involve the grouping of sentences together into higher level information blocks and their linking by means of extrinsic signals are essential.

It has already been mentioned that the practice of requiring students to write single paragraphs before tackling full-length texts has resulted in a tendency for second-language writers to treat paragraphs as isolated units. Thus students should also be taught how to provide signalling devices between adjacent paragraphs, signalling devices which will be macro-structural in function. Such higher level signals will often be metalingual devices in the form of whole clauses, sentences or even short paragraphs, an area usually ignored by the teachers. Because of their common structural pattern (see Chapter 7, Section 2.2) bridge sentences can be directly taught by presenting

the student with a limited number of formulae. Bridge paragraphs can also be produced using similar formulae.

Extrinsic signalling is an area where a limited teaching investment can produce high dividends, even with weak students. Of course, it should not be forgotten that signals must indicate something. They may aid the comprehension of a poorly written piece of discourse, but if the writing is totally lacking in structure no amount of signalling will suffice.

5. Some more general implications

Particularly at tertiary level, there should be very much more talking about language than there is at present, though it is important that this is always combined with the practice of language. Personal experience suggests that unlike grammar, rhetoric can be effectively taught as a content subject, and carry-over will be achieved if sufficient related practical writing takes place. This is, of course, provided that students are never assumed to possess the intuitive rhetorical competence available to the majority of educated mother-tongue speakers, and that as a result prescriptions are made as explicit as possible.

The study of models should be an integral part of the writing course, and this suggests that reading comprehension and writing should not appear on the syllabus as separate skills. The analysis and discussion of the means by which the writer organises and presents his message is just as important for the student as the interpretation of the message itself, the more usual activity of the comprehension class.

These models should, however, as far as possible always be authentic. While the presentation of specially constructed passages to intermediate students of EFL or ESP can be justified, the advanced ESL learner is a different case. He has already been exposed to a certain amount of authentic written language both inside and outside the classroom, and should ideally be aiming to develop the mother tongue writer's intuitive ability over a range of registers and styles. It is important for him to realise that at the supra-sentential level language is not <u>rigidly</u> rule-bound : that individual writers have slightly differing solutions to the problem of communication.

Finally, it is important that the teacher of English should have a clear overall picture of expository prose structure. This seems a trite and

obvious statement, but too often in developing countries where English is used as a second language the average teacher, even at secondary school level, has little writing ability, and certainly no theoretical knowledge of discourse structure. The necessity for rhetorically based writing courses combining both theory and practice in teacher training colleges can be seen when one considers the comments familiar to every teacher which continually appear on students' essays, such comments as, 'Muddled', 'Loosely written', 'Disconnected ideas', or 'Poorly organised', among others. Remarks like this reflect the gut reaction of the teacher rather than provide a means of improvement for the student. The teacher is aware that all is not well with the presentation of the subject-matter, that it does not read fluently, its ideas do not 'flow', but is usually not able to pinpoint what has gone wrong. A clearer conception of text organisation on the part of the teacher would enable him to point out more exactly where the student has erred, and suggest a practical pedagogical solution. A major object of this work, therefore, has been to provide a basis from which the teacher might be given this clearer conception, and thus be more able to deal with such problems adequately and explicitly.

6. An area for future research : process rather than product

Such a work is of necessity limited, no suggestion is being made that this is the only way of looking at text : indeed, the model presented here reflects just one of many possible 'realities', each 'reality' with its own pedagogical contribution to make. For example, this work has been exclusively product-orientated, but writing can equally well be viewed as a process. It is necessary to know as much as possible about the product before we can ascertain what the student should be aiming at, but the second stage is to discover the most efficient means of reaching this goal.

Almost nothing is known about the strategies employed by the individual writer. We all recognise those we ourselves apply, but there is no reason for supposing that there are not many possible routes to the same destination. However, it is conceivable that there is a limited number of most efficient strategies, and if some of these could be identified, each might be linked with the type of learner/writer for whom it promises the most success.

A closely associated area at present almost completely ignored is that of the relationship between expository writing and creative thinking. The

tendency to view writing simply as a product inhibits the teacher from helping the student to organise, develop and create his own ideas while actually engaged in the activity of writing. Some students express positive dissatisfaction about producing a detailed written plan before beginning to write; it is conceivable that some of these students are examples of the type of writer whose plan is created by the writing itself, who most actively uses the process as an aid in the organisation and development of his own thought.

Such areas are difficult areas to deal with experimentally, but the possible results could have a greater impact on the teaching of writing skills than any more traditional emphasis on product alone.

BIBLIOGRAPHY

Abbot, G. (1978). Motivation, materials, manpower and methods : some fundamental problems in ESP in ELT Documents 103 - Individualisation in language learning. The British Council : ETIC, 98-104.

Allbutt, T.C. (1904). Notes on the composition of scientific papers. London : Macmillan.

Bach, E. & R.T. Harms (eds). (1968). Universals in linguistic theory. Holt, Rinehart & Winston.

Bain, A. (1866). English composition and rhetoric. London : Longman, Green & Co.

Baldwin, C.S. (1902). A college manual of rhetoric. London : Longman, Green & Co.

Becker, A.L. (1965). A tagmemic approach to paragraph analysis College Composition and Communication 16, 5 (December) : 237-242. Reprinted in Symposium on the paragraph (1966).

Bhatia, A.T. (1974). An error analysis of students' compositions IRAL Vol. 12, 4

Biddulph, G.M.R. (1971). English studies series 11 : geography. London : Oxford University Press.

Bloomfield, L. (1933). Language New York : Holt, Rinehart & Winston.

Borissavlietch, M. (1958). The golden number. London Tiranti.

Bransford, J.D. & J.J. Franks (1972). The abstraction of linguistic ideas : a review. Cognition 1 : 211-250.

Bransford, J.D. & M.K. Johnson (1972). Contextual pre-requisites for understanding : some investigations of comprehension and recall. Journal of Verbal Learning and Verbal Behaviour, 11, 717-726.

Britton, J. et al (1975). The development of writing abilities (11-18). Schools Council Research Studies, MacMillan Education.

Brooks, C. & R.P. Warren (1950). Fundamentals of good writing : A handbook of modern rhetoric. New York : Harcourt, Brace & Co. Inc.

Canby, H.S. et al (1909). English composition in theory and practice. New York : Macmillan.

Chafe, W. (1974). Language and consciousness. Language 50, 1 111-133.

Chaplen, F. (1970). Paragraph writing. London, Oxford University Press.

Christensen, F. (1963). A generative rhetoric of the sentence College Composition and Communication 14, 3 (October) 155-161. Reprinted in Christensen (1967).

Christensen, F. (1965). A generative rhetoric of the paragraph. College Composition and Communication 16, 3 (October). Reprinted in Christensen (1967).

Christensen, F. (1967). Notes towards a new rhetoric. New York : Harper & Row.

Christensen, F. (1968). The problem of defining a mature style. English Journal 572-579.

Cochran, W.J. (1954). Some methods for strengthening the common X^2 tests. Biometrics 10, 417-451.

Connors, B. (1974). Writing a report. Milton Keynes : Open University Press (Social Science Third Level Course, Public Administration, Block 3 Part 3).

Corbett, E.P.J. (1971). Classical rhetoric for the modern student. New York : Oxford University Press.

Crothers, E.J. (1972). Memory structure and the recall of discourse. In Freedle & Carroll (eds). 247-284.

Coulthard, M. (1977). An introduction to discourse analysis. London : Longman.

Danes, F. (1974). Functional sentence perspective and the organisation of the text. In Danes, F. (ed). 106-127.

Danes, F. (1974). Papers on functional sentence perspective. The Hague : Mouton.

Dauterman, F.P. (1969). The syntactic structures employed in samples of narrative writing by secondary school students. Ph.D. Thesis, Ohio State University.

van Dijk, T.A. (1977a). Semantic macro-structures and knowledge frames in discourse comprehension. In Just & Carpenter, 3-32.

van Dijk, T.A. (1977b). Text and context : explorations in the semantics and pragmatics of discourse. London : Longman.

Donley, M. (1976). The paragraph in advanced composition : a heuristic approach. English Language Teaching Journal 30. 224-234.

Dressler, W.U. (1978). Current trends in textlinguistics. Berlin & New York : Walter de Gruyter.

E.L.T. Profile : Malawi (July 1975). British Council (mimeo).

Emig, J.A., J.T. Fleming & H.L. Popp (1966). Language and learning. New York : Harcourt, Brace Inc.

Erskine, J. (1946). The craft of writing. In Twentieth Century English Philosophical Library.

Everett, B.S. (1977). The analysis of contingency tables. London : Chapman and Hall.

Examination in English for the Malawi Certificate of Education : teaching programme, forms 3 and 4 (1973) Blantyre, Ministry of Education (mimeo, Ref. B139/2).

Fillmore, C.J. (1968). The case for case. In Bach & Harms, 1-88.

Flesch, R. (1949). The art of readable writing. New York : Harper.

Fowler, H.W. & F.G. Fowler (1930). The King's English. Oxford : The Clarendon Press.

Fredericksen, C.H. (1972). Effects of task-induced cognitive operations on comprehension and memory processes. In Freedle & Carroll, 211-246.

Freedle, R.O. & J. Carroll (eds) (1972). Language comprehension and the acquisition of knowledge. Washington DC : V.C. Winston & Sons.

Freeman, D.C. (ed) (1970). Linguistics and literary style. New York : Holt, Rinehart & Winston.

Fries, C.C. (1952). The structure of English. London : Longmans Green & Co.

Genung, J.F. (1886). The practical elements of rhetoric with illustrative examples. Boston : Ginn & Co.

Genung, J.F. & C.L. Hanson (1915). Outlines of composition and rhetoric. Boston : Ginn & Co.

Goodman, K.S. (1967). Reading : a psycholinguistic guessing-game. Journal of the Reading Specialist. May, 259-271.

Gowers, Sir E. (1973). The complete plain words (rev. B. Fraser). Harmondsworth : Penguin Books Ltd.

Gray, B. (1977). From discourse to dialog. Journal of Pragmatics 1, 283-298.

Grimes, J. (1972). Outlines and overlays. Language 48 513-524.

Grimes, J. (1975). The thread of discourse. The Hague : Mouton.

Grimes, J. (1978). Narrative studies in oral texts. In Dressler, 123-129.

Gunning, R. (1968). The technique of clear writing (Revised edition). McGraw Hill.

Haggett, P. (1975). Geography : a modern synthesis. New York : Harper & Row (International Edition).

Halliday, M.A.K. (1967). Notes on transitivity and theme in English Part 2. Journal of Linguistics 3, 199-244.

Halliday, M.A.K. & R. Hasan (1976). Cohesion in English London : Longman.

Hardie, J.L. (1938). The craft of composition Book 1 : the elements of composition. London : George Harrap & Co. Ltd.

Harris, Z. (1952). Discourse analysis. Language 28, 1-30.

Hausenblas, K. (1969). Krátká úvaha no téma "téma". (A short paper on the theme "theme"). CL. 17, 3-10.

Horn, V. (1968). An analysis of certain structural and lexical signals used to express relationships between sentences in paragraphs. EdD Thesis, University of Columbia.

Hulbert, J.R. & V.B. Hulbert (1929). Effective English. Chicago : University of Chicago Press.

Hunt, K.W. (1965). Grammatical structures written at three grade levels. N.C.T.E. Research Report 3, Champaign, Illinois.

Hunt, K.W. (1970a). Syntactic maturity in school-children and adults. Monographs of the Society for Research in Child Development, 35, 1, University of Chicago Press.

Hunt, K.W. (1970b). Recent measures in syntactic development. In Lester (ed). 187-200.

Hunt, K.W. (1971). Teaching syntactic maturity. In Perren & Trim, 287-300.

Imhoof, M. and H. Hudson (1975). From paragraph to essay. London : Longman.

Irmscher, W.F. (1976). The Holt guide to English : a contemporary handbook of rhetoric, language and literature. New York : Holt, Rinehart & Winston.

J.C.E. Syllabus : English (1973). Blantyre : Ministry of Education (mimeo).

Jones, A.E. & C.W. Faulkner (1951). Writing good prose : a simple structural approach. New York : Charles Scribner's Sons.

Jones, L.K. (1977). Theme in English expository discourse. Lake Bluff : Illinois Jupiter Press.

Jupp, T. & J. Milne (1972). Guided paragraph writing. London : Heinemann.

Just, M.A. & P.A. Carpenter (1977). Cognitive processes in comprehension. Hillsdale, N.J. : Lawrence Erlbaum Associates.

Kierzek, J.M. (1939). The MacMillan handbook of English. New York : MacMillan.

Kane, T.S. & L.J. Peters (1966). A practical rhetoric of expository prose. New York : Oxford University Press.

Kaplan, R.B. (1970). Notes towards an applied rhetoric. in Lugton, R.C. (ed) (b) 45-73.

Kaplan, R.B. (1972). The anatomy of rhetoric : prolegomena to a functional theory of rhetoric. The Centre for Curriculum Development Inc., Philadelphia, Pa.

Karrfalt, D.H. (1968). The generation of paragraphs and larger units. College Composition and Communication. 19, 3 211-217.

Katz, J.J. & J.A. Fodor (1963). The structure of a semantic theory. Language 39, 170-210.

Kieras, D.E. (1978). Good and bad structure in simple paragraphs : effects on apparent theme, reading time and recall. Journal of Verbal Learning and Verbal Behaviour 17 13-28.

Klare, G.R. (1974). Assessing readability. Reading Research Quarterly 10, 1 62-102.

Knapp, D.S. (1967). Formal factors affecting paragraph division in expository writing. Ed D Thesis, University of Columbia.

Koen, F.M., A.L. Becker & R.E. Young (1968). The psychological reality of the paragraph. In Zale, E.M. (ed). 174-187.

Lackstrom, J.E., L. Selinker & L.P. Trimble (1970). Grammar and technical English. In Lugton, R.C. (ed) (1970b) 101-133.

Lackstrom, J.E., L. Selinker & L.P. Trimble (1973). Technical rhetorical principles and grammatical choice. TESOL Quarterly 7, 2, 127-136.

Larson, R.L. (1967). Sentences in action : a technique for analysing paragraphs. College Composition and Communication 13, February, 16-22.

Leech, G. (1974). Semantics. Harmondsworth : Penguin.

Lester, M. (ed) (1970). Readings in applied transformational grammar. New York : Holt, Rinehart & Winston Inc.

Lewis, E.H. (1894). The history of the English paragraph. Chicago : University of Chicago Press.

Lewontin, R.C. & J. Felsenstein (1965). The robustness of homogeneity tests in 2XN tables. Biometrics 21, 19-33.

Litteral, R. (1972). Rhetorical predicates and time typology in Anggor. Foundations of Language 8, 391-410.

Lugton, R.C. (ed) (1970a). Preparing the EFL teacher : a projection for the seventies. Centre for Curriculum Development Inc. : Philadelphia, P.A.

Lugton, R.C. (ed) (1970b). English as a second language : current issues. Centre for Curriculum Development, Inc : Philadelphia, P.A.

Lyons, J. (1977). Semantics, Vol. 1. Cambridge : Cambridge University Press.

Mackay, W.F. (1965). Language teaching analysis. London : Longman.

M.C.E. Chief Examiner's Report,English, Paper 2. (1972). Blantyre : Ministry of Education. (mimeo).

M.C.E. Chief Examiner's Report, English, Paper 2. (1973). Blantyre : Ministry of Education. (mimeo).

M.C.E. Examination Board : Examination in English. (1972). Blantyre : Ministry of Education. (mimeo, Ref. B139/6 revised).

Mellon, J.C. (1967). Transformational sentence-combining : a method for enhancing the development of syntactic fluency in English composition. Report to the U.S. Office of Education Cooperative Research Project No 5-B418.

Meyer, B. (1975). The organisation of prose and its effects on memory. Amsterdam : North Holland Publishing Co.

Minsky, M. (1975). A framework for representing knowledge. In Winston (ed). 211-278.

Moulton, W.G. (1966). A linguistic guide to language teaching. Modern Languages Association of America.

Partridge, E. (1954). The concise usage and abusage : A modern guide to good English. London : Hamish Hamilton.

Perren, G.E. & J.L.M. Trim (eds) (1971). Applications of linguistics. Cambridge : Cambridge University Press.

Pike, K.L. (1954-60). Language in relation to a unified theory of human behaviour. The Hague : Mouton.

Pike, K.L. (1959). Language as particle, wave and field. The Texas Quarterly 2, 2 37-54.

Pike, K.L. & E.G. Pike (1977). Grammatical analysis. Summer Institute of Linguistics/University of Texas at Arlington.

Pitkins, W. (1969). Discourse blocs. College Composition and Communication 20, May, 2, 138-148.

Quiller-Couch, Sir A. (1916). On the art of writing : Lectures delivered to the University of Cambridge 1913-14. Cambridge : University of Cambridge Press.

Reading and thinking in English Vol 3 : Discovering discourse (1979). Oxford : Oxford University Press.

Rees, M. & A.H. Urquhart (1976). Intonation as a guide to readers' structuring of prose texts. Work in Progress 9, University of Edinburgh Department of Linguistics.

Rivers, W.A. (1964). The psychologist and the foreign language teacher. Chicago : University of Chicago Press.

Rodgers, P.C. (1965). Alexander Bain and the rise of the organic paragraph. Quarterly Journal of Speech 51, December 399-408.

Rodgers, P.C. (1966a). A discourse-centred rhetoric of the paragraph. College Composition and Communication 17, 1, February, 2-11.

Rodgers, P.C. (1966b). A reply to Christensen and Becker. College Composition and Communication 17, 2, May.

Rodgers, P.C. (1966c). Comment in Symposium on the paragraph.

Schlesinger, I.M. (1968). Sentence structure and the reading process. The Hague : Mouton.

The sentence and the paragraph (1966). National Council for Teachers of English Urbana : Illinois.

Shaw, H. (1946). Writing and re-writing : Books 1 and 2 of a complete course in freshman English. New York : Harper & Bros.

Sherman, L.A. (1893). Analytics of literature : a manual for the objective study of English prose and poetry. Boston : Ginn & Co.

Slakter, M.J. (1966). Comparative validity of the chi-square and two modified chi-square goodness of fit tests for small but equal expected frequencies. Biometriks 53, 619-623.

Smart, W.K. (1922). Handbook of effective writing. New York : Harper & Bros.

Stamp, L.D. (1960). Our developing world. London : Faber & Faber.

Symposium on the paragraph (1966). College Composition and Communication 17, 2 (May) Reprinted in The Sentence and the paragraph.

Thomson, T. (1948). An outline of forestry. London : Allen & Unwin Ltd.

Thompson, W.N. (1957). Fundamentals of communication : an integrated approach. New York : McGraw Hill Book Co Inc.

Tufte, V. (1971). Grammar as style. New York : Holt, Rinehart & Winston.

Urquhart, A.H. & H.G. Widdowson (1976). K.A.A.U. English for academic purposes project. 1st Yearly Report : K.A.A.U. Reading Research Project, University of Edinburgh.

Vallins, G.H. (1951). Good English : how to write it. London : Pan Books Ltd.

Watson, L. (1973). Supernature. London : Hodder & Stoughton.

Wendell, B. (1893). English composition : eight lectures given at the Lowell Institute. New York.

Wehrlich, E. (1976). A text grammar of English. Heidelberg : Quelle and Meyer.

Whitten, W. & F. Whitaker (1939). Good and bad English : a guide to speaking and writing. London : George Newnes, Ltd.

Widdowson, H.G. (1973). An applied linguistic approach to discourse analysis. Ph.D. Thesis, University of Edinburgh.

Widdowson, H.G. (1979). Explorations in applied linguistics. Oxford : Oxford University Press.

Wilkins, D.A. (1972). Linguistics in language teaching. London : Arnold.

Winograd, T. (1977). A framework for understanding discourse. In Just & Carpenter (eds.) 63-88.

Winston, P.H. (ed) (1975). The psychology of computer vision. New York : McGraw Hill

Wohl, M. (1978). Techniques for writing composition. Rowley, Mass : Newbury House Publishers, Inc.

Young, R.E. & A.L. Becker (1964). The role of lexical and grammatical cues in paragraph recognition. Studies in Language and Language Behaviour 1-6. Centre for Research in Language and Language Behaviour, Ann Arbor : Michigan University Press.

Young, R.E. & A.L. Becker (1966). Towards a modern theory of rhetoric : a tagmemic contribution. In Emig, Fleming & Popp (eds), 195-215.

Zale, E.M. (ed) (1968). Proceedings of the Conference on Language and Language Behaviour. New York : Appleton-Century-Crofts.

A P P E N D I X 1

FIFTEEN REPRESENTATIVE TEXTBOOKS OF TRADITIONAL RHETORIC

1866 - 1978

APPENDIX 1

REPRESENTATIVE TEXTBOOKS OF TRADITIONAL RHETORIC, 1866-1978

1. Bain, Alexander (1866) English Composition and Rhetoric London : MacMillan.

2. Genung, John F. (1886) The Practical Elements of Rhetoric with Illustrative Examples Boston : Ginn & Co.

3. Baldwin, Charles Sears (1902) A College Manual of Rhetoric London : Longman : Longman Green & Co.

4. Canby, Henry Seidel et al (1909) English Composition in Theory and Practice New York : MacMillan.

5. Scott, F. N. & J. V. Denney (1911) The New Composition-Rhetoric Boston & Chicago : Allyn & Bacon.

6. Genung, John F. & C. L. Hansen (1915) Outlines of Composition and Rhetoric Boston : Ginn & Co.

7. Smart, W. K. (1922) Handbook of Effective Writing New York : Harper & Bros.

8. Hulbert, J. R. & V. B. Hulbert (1929) Effective English Chicago, Illinois : University of Chicago Press.

9. Kierzek, John M. (1939) The MacMillan Handbook of English New York : MacMillan.

10. Shaw, Harry (1946) Writing and Re-Writing : Books 1 and 2 of a Complete Course in Freshman English New York : Harper & Bros.

11. Brooks, Cleanth & Robert Penn Warren (1950) Fundamentals of Good Writing : A Handbook of Modern Rhetoric New York : Harcourt, Brace & Co Inc.

12. Thompson, Wayne N. (1957) Fundamentals of Communication : an Integrated Approach New York : McGraw Book Co Inc.

13. Jones, A. E. & C. W. Faulkner (1961) Writing Good Prose : A Simple Structural Approach New York : Charles Scribner's Sons.

14. Kane, Thomas & Leonard J. Peters (1966) A Practical Rhetoric of Expository Prose New York : Oxford University Press.

15. Wohl, Milton (1978) Techniques for Writing : Composition Rowley, Mass. : Newbury House Publishers Inc.

APPENDIX 2a

PASSAGES USED FOR EXPERIMENT 1

(a) Medium sentence version

(b) Long sentence version

(c) Varied sentence version

(d) Short sentence version

APPENDIX 2a

EXPERIMENT 1 : EARTHQUAKES PASSAGE

(a) Earthquakes and volcanoes are both serious natural hazards which cause millions of pounds worth of damage every year. They take the lives of thousands of people, and even those who escape are often left homeless. Rescue-work and repairs are costly and time-consuming, often necessitating international cooperation. Earthquakes are particularly disastrous to man because he does not know when or where they will occur, and they may cause whole towns and villages to collapse. Blocks of flats are badly damaged, houses fall down, water-pipes burst, electricity cables are broken, and railway tracks are torn up. Fires break out which are difficult to extinguish because there is no water supply, and many people may be killed or injured. Volcanoes are just as dangerous as earthquakes, volcanic activity also sometimes destroying whole towns, by burying them in lava, dust and ashes. The destruction of the Roman town of Pompeii in AD79 was the most famous natural catastrophe in history, and we can still see the remains to this very day. Man can do nothing to prevent such disasters, but some volcanoes show signs of activity for several days or even weeks before erupting seriously. Nowadays observatories are sometimes built near these volcanoes, so that scientists can watch them and forecast likely eruptions. The people living nearby may then be able to move away before disaster overtakes them.

(b) Earthquakes and volcanoes are both serious natural hazards which cause millions of pounds worth of damage every year and take the lives of thousands of people, even those who escape often being left homeless, so that rescue-work and repairs are costly and time-consuming, often necessitating international cooperation. Earthquakes are particularly disastrous to man because he does not know when or where they will occur, and they may cause whole towns or villages to collapse, so that blocks of flats are badly damaged, houses fall down, water-pipes burst, electricity cables are broken, railway tracks are torn up, fires break out which are difficult to extinguish because there is no water supply, and many people may be killed or injured. Volcanoes are just as dangerous as earthquakes, volcanic activity also sometimes destroying whole towns, by burying them in lava, dust and ashes, for example, the destruction of the Roman town of Pompeii in AD79, the most famous natural catastrophe in history, the remains of which we can still see to this very day. Although man can do nothing to prevent such disasters, some volcanoes show signs of activity for several days or even weeks before erupting seriously, and nowadays observatories are sometimes built near these volcanoes for scientists to watch them and forecast likely eruptions, with the result that the people living nearby may be able to move away before disaster overtakes them.

(c) Earthquakes and volcanoes are both serious natural hazards which cause millions of pounds worth of damage every year and take the lives of thousands of people. Even those who escape are often left homeless, and rescue-work and repairs are costly and time-consuming, often necessitating international cooperation. Earthquakes are particularly disastrous to man. He does not know when or where they will occur, and they may cause whole towns or villages to collapse, so that blocks of flats are badly damaged, houses fall down, water-pipes burst, electricity cables are broken, and railway tracks are torn up. Fires break out which are difficult to extinguish because there is no water supply and many people may be killed or injured. Volcanoes are just as dangerous as earthquakes. Volcanic activity sometimes destroys whole towns by burying them in lava, dust and ashes, as happened at the destruction of the Roman town of Pompeii in AD79. This was the most famous natural catastrophe in history, and we can still see its remains to this very day. Man can do nothing to prevent such disasters. Some volcanoes, however, show signs of activity for several days or even weeks before erupting seriously, and nowadays observatories are sometimes built near these volcanoes. Scientists can then watch them and forecast likely eruptions, so that the people living nearby may be able to move away before disaster overtakes them.

(d) Earthquakes and volcanoes are both serious natural hazards. They cause millions of pounds worth of damage every year. They take the lives of thousands of people. Those who escape are often left homeless. Rescue-work and repairs are costly and time-consuming. They often necessitate international cooperation. Earthquakes are particularly disastrous to man. He does not know when or where they will occur. They may cause whole towns or villages to collapse. Blocks of flats are badly damaged. Houses fall down. Water-pipes burst. Electricity cables are broken. Railway tracks are torn up. Fires break out. They are difficult to extinguish. This is because there is no water supply. Many people may be killed or injured. Volcanoes are just as dangerous as earthquakes. Volcanic activity also sometimes destroys whole towns. It may bury them in lava, dust and ashes. The Roman town of Pompeii was destroyed in AD79. This was the most famous natural catastrophe in history. We can still see its remains to this very day. Man can do nothing to prevent such disasters. Some volcanoes, however, may show signs of activity before erupting seriously. This may occur for several days or even weeks. Nowadays observatories are sometimes built near these volcanoes. Scientists can watch them. They can then forecast likely eruptions. The people living nearby may be able to move away before disaster overtakes them.

1	2	3	4

APPENDIX 2a

EXPERIMENT 1 : SOILS PASSAGE

(a) The best texture for a good agricultural soil is that of a loam. This is a mixture of fine sand or silt with a little clay to hold the moisture and prevent it draining through the soil too rapidly. This mixture should not be stony, as stones prevent the proper development of root systems, and may damage the plough. From time immemorial the significance of soil texture has been recognised. The ancient Egyptians knew the virtues of Nile alluvium, waiting anxiously for the annual flood before planting their crops. The Anglo-Saxon settlers in Britain sought the loam terrains, avoiding the heavy clays and settling on the hills. Sometimes such soils do not exist, and then they have often been made, as, for example, in China. There, a farmer might live on heavy land while his neighbour farmed sandy land, so they would exchange a few thousand baskets of soil. Enough labour was available, and the mixture, once made, was more or less permanent. Similarly in Britain, where the early farmer lived on light superficial soils he would dig clay or marl from pits to spread over his land. Though the importance of texture is known, and powerful earth-moving machinery is available, no attempts are being made today to make good agricultural land. This could easily be done by mixing the correct proportions of sand, silt and clay together, so we should make the attempt, good agricultural land being vital to Britain's economy.

(b) The best texture for a good agricultural soil is that of a loam, a mixture of fine sand or silt with a little clay to hold the moisture and prevent it draining through the soil too rapidly, but this mixture should not be stony, as stones prevent the proper development of root systems and may damage the plough. From time immemorial the significance of soil texture has been recognised, for example by the ancient Egyptians, who knew the virtues of Nile alluvium, waiting anxiously for the annual flood before planting their crops, and the Anglo-Saxon settlers in Britain sought the loam terrains, avoiding the heavy clays and settling on the hills. If such soils do not exist they have often been made, as in China, where a farmer might live on heavy land while his neighbour lived on sandy land, so they would exchange a few thousand baskets of soil, enough labour being available, and the mixture, once made, being more or less permanent, or in Britain, where the early farmer living on light superficial soils would dig clay or marl from pits to spread over his land. The importance of texture is known, and powerful earth-moving machinery is available, but in spite of this no attempts are being made today to make good agricultural soil, although this could easily be done by mixing the correct proportions of sand, silt and clay together, and an attempt at it should be made, as good agricultural land is vital to Britain's economy.

(c) The best texture for a good agricultural soil is that of a loam. This is a mixture of fine sand or silt with a little clay to hold the moisture and prevent it draining through the soil too rapidly. This mixture should not be stony, as stones prevent the proper development of root systems and may damage the plough. From time immemorial the significance of soil texture has been recognised. For example, the ancient Egyptians knew the virtues of Nile alluvium, waiting anxiously for the annual flood before planting their crops, and the Anglo-Saxon settlers in Britain sought the loam terrains, avoiding the heavy clays and settling on the hills. Sometimes such soils do not exist. Then they have often been made, as in China, where, if a farmer lived on heavy land while his neighbour farmed sandy land, they would exchange a few thousand baskets of soil. Enough labour was available, and once the mixture was made it was more or less permanent. Similarly in Britain, where the early farmer lived on light superficial soils he would dig clay or marl from pits to spread over his land. The importance of texture is therefore known. Nevertheless, although powerful earth-moving machinery is available, no attempts are being made today to make good agricultural soil, although this could easily be done by mixing the correct proportions of sand, silt and clay together. An attempt at this should be made, however, as good agricultural land is vital to Britain's economy.

(d) The best texture for a good agricultural soil is that of a loam. This is a mixture of fine sand or silt with a little clay. The clay holds the moisture. It prevents it draining through the soil too rapidly. The mixture should not be stony. Stones prevent the proper development of root systems. They may also damage the plough. From time immemorial the significance of soil texture has been recognised. The ancient Egyptians knew the virtues of Nile alluvium. They waited anxiously for the annual flood. Then they planted their crops. The Anglo-Saxon settlers in Britain sought the loam terrains. They avoided the heavy clays. They settled on the hills. Sometimes such soils do not exist. Then they have often been made. In China a farmer might live on heavy land. His neighbour might farm sandy land. They would exchange a few thousand baskets of soil. Enough labour was available for this. Once the mixture was made, it was more or less permanent. In Britain the early farmer might live on light superficial soils. He would therefore dig clay or marl from pits. He would spread this over his land. The importance of textureis known. Powerful earth-moving machinery is available. Nevertheless, no attempts are being made today to make good agricultural soil. This could easily be done. The correct proportions of sand, silt and clay could be mixed together. An attempt at this should be made. Good agricultural land is vital to Britain's economy.

1	2	3	4

APPENDIX 2a

EXPERIMENT 1 : TORNADOES PASSAGE

(a) Tornadoes are the most violent and destructive of all atmospheric disturbances, but fortunately they are quite small, and the areas they affect are limited. A tornado appears as a writhing column of cloud, reaching down to earth, and apparently suspended from a thick dark cloud above. Within the column the air is whirling round at a very high speed. At the centre the air becomes rarified under the influence of the centrifugal force, and its pressure falls to half the normal value. The central area of heavy damage is only 100 yards wide, but as the tornado moves across the countryside it leaves a trail of complete destruction in its wake. Tornadoes occur at cold fronts when the warm air is especially damp and unstable. The worst tornadoes are found in the Middle-West of the United States, and to a lesser extent in Australia. Dust devils, in a sense miniature tornadoes, are whirling columns of air occurring in desert areas. They suck up quantities of sand, and result from atmospheric instability caused by intense heating. Dust storms occur in dry desert areas, when rising currents associated with cold fronts lift large quantities of dust into the air. They appear as large walls of dust advancing along the line of the front, the best-known being those of the Sudan, where they are called 'haboobs'.

(b) Tornadoes are the most violent and destructive of all atmospheric disturbances, but fortunately they are quite small, and the areas they affect are quite limited. A tornado appears as a writhing column of cloud, reaching down to earth, and apparently suspended from a thick dark cloud above, and within it the air is whirling around at a very high speed. At the centre of the tornado the air becomes rarified under the influence of the centrifugal force, with the air pressure falling to half the normal value, and this centre, only 100 yards wide, is the area of heavy damage, so that as the tornado moves across the countryside it leaves a trail of complete destruction in its wake. Tornadoes occur at cold fronts when the warm air is especially damp and unstable, the worst ones being found in the Middle-West of the United States and to a lesser extent in Australia. Dust devils, which in a sense are miniature tornadoes, are whirling columns of air occurring in desert areas, where they suck up quantities of sand as a result of atmospheric instability caused by intense heating, while dust storms occur in dry desert areas when rising currents associated with cold fronts lift large quantities of dust into the air, so that they appear as large walls of dust advancing along the line of the front, the best-known being those of the Sudan, where they are called 'haboobs'.

(c) Tornadoes are the most violent and destructive of all atmospheric disturbances, but fortunately they are quite small, and the areas they affect are limited. A tornado appears as a writhing column of cloud. This column, reaching down to earth, is apparently suspended from a thick dark cloud above, and within it the air is whirling around at a very high speed. At the centre the air becomes rarified under the influence of the centrifugal force, and its pressure falls to half the normal value. The central area of heavy damage is only 100 yards wide, but as the tornado moves across the countryside it leaves a trail of complete destruction in its wake. Tornadoes occur at cold fronts when the warm air is especially damp and unstable. The worst tornadoes are found in the Middle-West of the United States, and to a lesser extent in Australia. Dust devils are in a sense miniature tornadoes. They are whirling columns of air occurring in desert areas which suck up quantities of sand, and result from atmospheric instability caused by intense heating. Dust storms also occur in dry desert areas. These are produced when rising currents associated with cold fronts lift large quantities of dust into the air, and appear as large walls of dust advancing along the line of the front. The best-known dust storms are those of the Sudan, where they are called 'haboobs'.

(d) Tornadoes are the most violent and destructive of all atmospheric disturbances. Fortunately they are quite small. Also the areas they affect are limited. A tornado appears as a writhing column of cloud. It reaches down to earth. It is apparently suspended from a thick dark cloud above. Within the column the air is moving around. This is happening at a very high speed. At the centre the air becomes rarified. This is the result of the centrifugal force. The air pressure falls to half the normal value. The centre of the tornado is the area of heavy damage. This area is only 100 yards wide. The tornado moves across the countryside. It leaves a trail of complete destruction in its wake. Tornadoes occur at cold fronts. They occur when the warm air is especially damp and unstable. The worst ones are found in the Middle-West of the United States. They are also found to a lesser extent in Australia. Dust devils are in a sense miniature tornadoes. They are whirling columns of air. They occur in desert areas. They suck up quantities of sand. They result from atmospheric instability. This is caused by intense heating. Dust storms occur in dry desert areas. They appear when rising currents lift large quantities of dust into the air. These currents are associated with cold fronts. A dust storm appears as a large wall of dust. This wall is advancing along the line of the front. The best-known dust storms are those of the Sudan. There they are called 'haboobs'.

1	2	3	4

A P P E N D I X 2 b

PASSAGES USED FOR EXPERIMENT 2

(a) Short sentences at change of topic

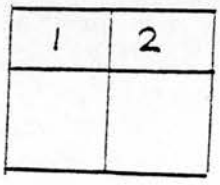
(b) Short sentences randomly placed

APPENDIX 2b

EXPERIMENT 2 : EARTHQUAKES PASSAGE

(a) Earthquakes and volcanoes are both serious natural hazards which cause millions of pounds worth of damage every year and take the lives of thousands of people. Even those who escape are often left homeless, and rescue work and repairs are costly and time-consuming, often necessitating international cooperation. Earthquakes are particularly disastrous to man. He does not know when or where they will occur, and they may cause whole towns or villages to collapse, so that blocks of flats are badly damaged, houses fall down, water-pipes burst, electricity cables are broken, and railway tracks are torn up. Fires break out which are difficult to extinguish because there is no water supply and many people may be killed or injured. Volcanoes are just as dangerous as earthquakes. Volcanic activity sometimes destroys whole towns by burying them in lava, dust and ashes, as happened at the destruction of the Roman town of Pompeii in AD79. This was the most famous natural catastrophe in history, and we can still see its remains to this very day. Man can do nothing to prevent such disasters. Some volcanoes, however, show signs of activity for several days or even weeks before erupting seriously, and nowadays observatories are sometimes built near these volcanoes. Scientists can then watch them and forecast likely eruptions, so that the people living nearby may be able to move away before disaster overtakes them.

(b) Earthquakes and volcanoes are both serious natural hazards which cause millions of pounds worth of damage every year and take the lives of thousands of people. Even those who escape are often left homeless, and rescue work and repairs are costly and time-consuming, often necessitating international cooperation. Earthquakes are particularly disastrous to man because he does not know when or where they will occur, and they may cause whole towns or villages to collapse, so that blocks of flats are badly damaged, houses fall down, water-pipes burst and electricity cables are broken. Railway tracks are torn up. Fires break out which are difficult to extinguish because there is no water supply, and many people may be killed or injured. Volcanoes are just as dangerous as earthquakes, volcanic activity sometimes destroying whole towns by burying them in lava, dust and ashes, as happened at the destruction of the Roman town of Pompeii. This was in AD79. It was the most famous natural catastrophe in history, and we can still see its remains to this very day. Man can do nothing to prevent such disasters, but some volcanoes show signs of activity for several days or even weeks before erupting seriously, and nowadays observatories are sometimes built near these volcanoes. Scientists can watch them. They can then forecast likely eruptions, so that the people living nearby may be able to move away before disaster overtakes them.

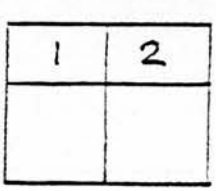


APPENDIX 2b

EXPERIMENT 2 : SOILS PASSAGE

(a) The best texture for a good agricultural soil is that of a loam. This is a mixture of fine sand or silt with a little clay to hold the moisture and prevent it draining through the soil too rapidly. This mixture should not be stony, as stones prevent the proper development of root systems and may damage the plough. From time immemorial the significance of soil texture has been recognised. For example, the ancient Egyptians knew the virtues of Nile alluvium, waiting anxiously for the annual flood before planting their crops, and the Anglo-Saxon settlers in Britain sought the loam terrains, avoiding the heavy clays and settling on the hills. Sometimes such soils do not exist. Then they have often been made, as in China, where, if a farmer lived on heavy land while his neighbour farmed sandy land, they would exchange a few thousand baskets of soil. Enough labour was available, and once the mixture was made it was more or less permanent. Similarly in Britain, where the early farmer lived on superficial soils he would dig clay or marl from pits to spread over his land. The importance of texture is therefore known. Nevertheless, although powerful earth-moving machinery is available, no attempts are being made today to make good agricultural soil, although this could easily be done by mixing the correct proportions of sand, silt and clay together. An attempt at this should be made, however, as good agricultural land is vital to Britain's economy.

(b) The best texture for a good agricultural soil is that of a loam, which is a mixture of fine sand or silt with a little clay to hold the moisture and prevent it draining through the soil too rapidly. This mixture should not be stony, as stones prevent the proper development of root systems and may damage the plough. From time immemorial the significance of soil texture has been recognised, for example by the ancient Egyptians, who knew the virtues of Nile alluvium and waited anxiously for the annual flood. Then they planted their crops. The Anglo-Saxon settlers in Britain sought the loam terrains, avoiding the heavy clay and settling on the hills. If such soils do not exist they have often been made, as, for example, in China, where a farmer might live on heavy land. His neighbour might farm sandy land. They would therefore exchange a few thousand baskets of soil, as enough labour was available, and once the mixture was made it was more or less permanent. Similarly in Britain, where the early farmer lived on light superficial soils he would dig clay or marl from pits to spread over his land. Although the importance of texture is known and powerful earth-moving machinery is available, no attempts are being made today to make good agricultural soil. This could easily be done. Sand, silt and clay could be mixed together in the correct proportions, and an attempt at this should be made, good agricultural land being vital to Britain's economy.

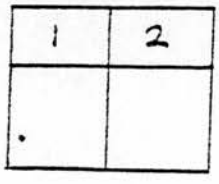


APPENDIX 2b

EXPERIMENT 2 : TORNADOES PASSAGE

(a) Tornadoes are the most violent and destructive of all atmospheric disturbances, but fortunately they are quite small and the areas they affect are limited. A tornado appears as a writhing column of cloud. This column, reaching down to earth, is apparently suspended from a thick dark cloud above, and within it the air is whirling around at a very high speed. At the centre the air becomes rarified under the influence of the centrifugal force, and its pressure falls to half the normal value. The central area of heavy damage is only 100 yards wide, but as the tornado moves across the countryside it leaves a trail of complete destruction in its wake. Tornadoes occur at cold fronts when the warm air is especially damp and unstable. The worst tornadoes are found in the Middle-West of the United States, and to a lesser extent in Australia. Dust devils are in a sense miniature tornadoes. They are whirling columns of air occurring in desert areas which suck up quantities of sand, and result from atmospheric instability caused by intense heating. Dust storms also occur in dry desert areas. These are produced when rising currents associated with cold fronts lift large quantities of dust into the air, and appear as large walls of dust advancing along the line of the front. The best-known dust storms are those of the Sudan, where they are called 'haboobs'.

(b) Tornadoes are the most violent and destructive of all atmospheric disturbances, but fortunately they are quite small, and the areas they affect are limited. A tornado appears as a writhing column of cloud, apparently suspended from a thick dark cloud above. It reaches down to earth. Within the column the air is whirling round at a very high speed, and at the centre it becomes rarified under the influence of the centrifugal force, its pressure falling to half the normal value, so that the centre is the area of heavy damage. It is only 100 yards wide. Nevertheless, as the tornado moves across the countryside it leaves a trail of complete destruction in its wake. Tornadoes occur at cold fronts when the warm air is especially damp and unstable, the worst being found in the Middle-West of the United States and to a lesser extent in Australia. Dust devils, in a sense miniature tornadoes, are whirling columns of air occurring in desert areas. They suck up quantities of sand. They are the result of instability in the atmosphere which has been caused by the intensity of the heating. Dust storms occur in dry desert areas, when rising currents associated with cold fronts lift large quantities of dust into the air. They appear as large walls of dust advancing along the line of the front, the best-known being those of the Sudan, where they are called 'haboobs'.



A P P E N D I X 2 c

MODIFIED KERNELS

(a) Earthquakes Passage

(b) Soil Passage

(c) Tornadoes Passage

APPENDIX 2c

(a) MODIFIED KERNELS : EARTHQUAKES

1. Earthquakes and volcanoes are both serious natural hazards.
2. They cause millions of pounds worth of damage every year.
3. They take the lives of thousands of people.
4. Those who escape are often left homeless.
5. Rescue work and repairs are costly and time-consuming.
6. They often necessitate international cooperation.
7. Earthquakes are particularly disastrous to man.
8. He does not know when or where they will occur.
9. They may cause whole towns and villages to collapse.
10. Blocks of flats are badly damaged.
11. Houses fall down.
12. Water-pipes burst.
13. Electricity cables are broken.
14. Railway tracks are torn up.
15. Fires often break out.
16. They are difficult to extinguish.
17. There is no water-supply.
18. Many people may be killed or injured.
19. Volcanoes are as dangerous as earthquakes.
20. Violent volcanic activity often destroys whole towns.
21. It may bury them in lava, dust and ashes.
22. The Roman town of Pompeii was destroyed.
23. This happened in AD 79.
24. It was the most famous natural catastrophe in history.
25. We can still see the remains to this very day.
26. Man can do nothing to prevent such disasters.
27. Some volcanoes may show signs of activity for several days or even weeks.
28. (Before) they erupt seriously.
29. Nowadays observatories are sometimes built near these volcanoes.
30. Scientists can watch them.
31. They can forecast likely eruptions.
32. They people living nearby may be able to move away.
33. (Before) disaster overtakes them.

(b) MODIFIED KERNELS : SOILS

1. The best texture for a good agricultural soil is that of a loam.
2. This is a mixture of fine sand or silt.
3. There is also a little clay.
4. The clay holds moisture.
5. It prevents the moisture draining through the soil too rapidly.
6. The mixture should not be stony.
7. Stones prevent the proper development of root systems.
8. Stones may damage the plough.
9. From time immemorial the significance of soil texture has been recognised.
10. The ancient Egyptians knew the virtues of Nile alluvium.
11. They waited anxiously for the annual flood.
12. (Then) they planted their crops.
13. The Anglo-Saxon settlers in Britain sought the loam terrains.
14. They avoided the heavy clays.
15. They settled on the hills.

16. Sometimes such soils do not exist.
17. (Then) they have often been made.
18. In China a farmer might live on heavy land.
19. His neighbour might farm sandy land.
20. They would exchange a few thousand baskets of soil.
21. Enough labour was available.
22. Once the mixture was made it was more or less permanent.
23. In Britain the early farmer might live on light, superficial soils.
24. He would dig clay or marl from pits.
25. He would spread this over his land.
26. The importance of texture is known.
27. Powerful earth-moving machinery is available.
28. No attempts are being made today to make good agricultural soil.
29. This could easily be done.
30. (This could be done) by moving the correct proportions of sand, silt and clay.
31. An attempt should be made.
32. Good agricultural land is vital to Britain's economy.

(c) MODIFIED KERNELS : TORNADOES

1. Tornadoes are the most violent and destructive of atmospheric disturbances.
2. Fortunately they are quite small.
3. (Fortunately) the areas they affect are limited.
4. A tornado appears as a writhing column of cloud.
5. This column reaches down to earth.
6. It is apparently suspended from a thick dark cloud above.
7. Within the column the air is whirling around.
8. (It is whirling around) at a very high speed.
9. At the centre the air becomes rarified.
10. This happens under the influence of the centrifugal force.
11. The air pressure falls to half the normal value.
12. The centre of the tornado is the area of heavy damage.
13. This area is only a hundred yards across.
14. The tornado moves across the countryside.
15. It leaves a trail of complete destruction in its wake.
16. Tornadoes occur at cold fronts.
17. This happens when the warm air is especially damp and unstable.
18. The worst tornadoes are found in the middle-west of the United States.
19. They are also found to a lesser extent in Australia.
20. Dust-devils are in a sense miniature tornadoes.
21. They are whirling columns of air.
22. They occur in desert areas.
23. They suck up quantities of sand.
24. They result from atmospheric instability.
25. This is caused by intense heating.
26. Dust-storms occur in dry desert areas.
27. (When) rising currents lift large quantities of dust into the air.
28. These currents are associated with cold fronts.
29. The appearance is of a large wall of dust.
30. It is advancing along the line of the front.
31. The best-known dust storms are those of the Sudan.
32. They are called there 'haboobs'.

APPENDIX 2d

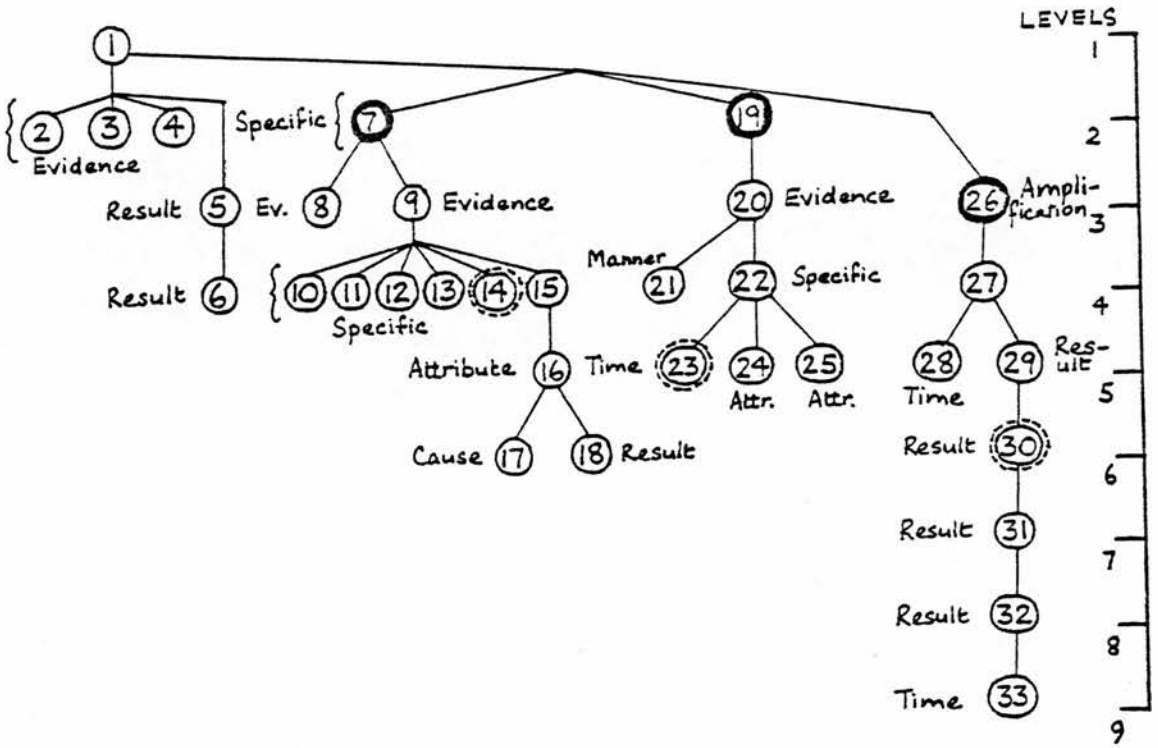
EXPERIMENT 2 : CONTENT STRUCTURE OF THE THREE PASSAGES

(a) Earthquakes Passage

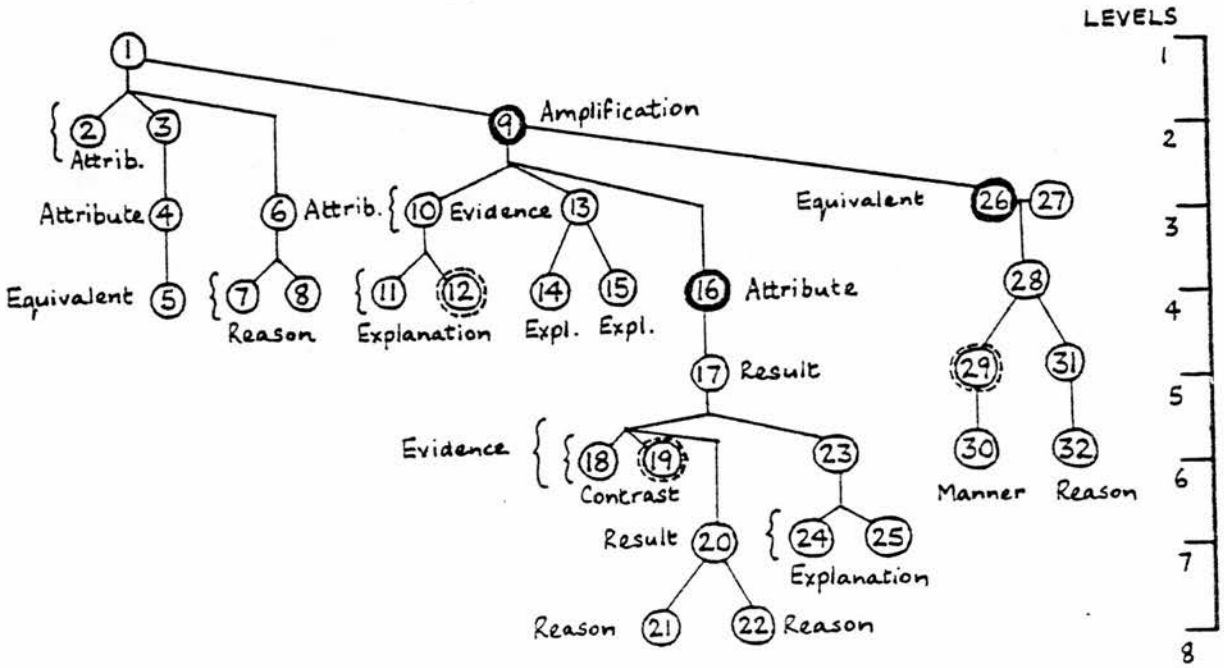
(b) Soil Passage

(c) Tornadoes Passage

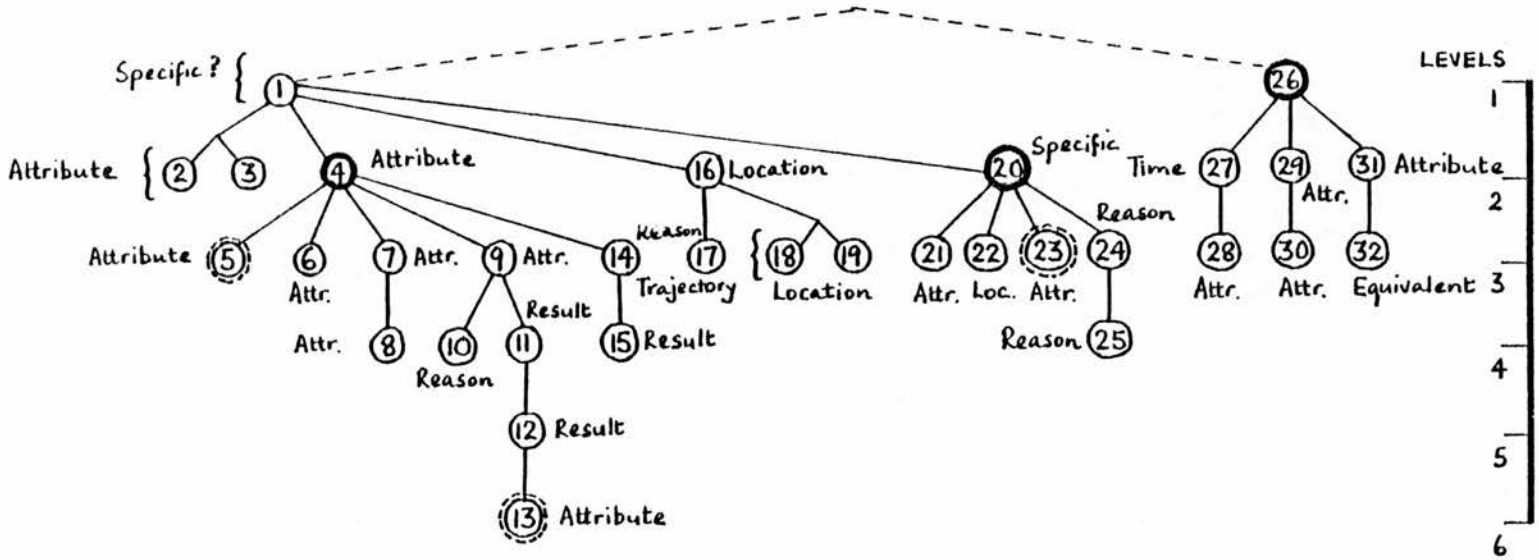
THE THREE PASSAGES : CONTENT STRUCTURE



(a) Earthquakes Passage



(b) Soil Passage



(c) Tornadoes Passage

KEY

O Short sentence at change of topic

(O) Short sentence randomly placed

A P P E N D I X 3

EXPERIMENT 1 : STATISTICAL ANALYSES

(a) Kendall's Coefficient of Concordance (W)

(b) Analysis of Residuals

(c) Somers's 'd'

(a) KENDALL'S COEFFICIENT OF CONCORDANCE (W)

	EARTHQUAKES				SOILS				TORNADOES			
N	A	B	C	D	A	B	C	D	A	B	C	D
1	2	1	4	3	1	2	3	4	2	1	4	3
2	2	3	4	1	1	2	4	3	1	3	2	4
3	3	1	4	2	2	1	3	4	1	2	3	4
4	2	1	4	3	4	1	3	2	3	1	2	4
5	2	1	4	3	1	2	4	3	2	1	3	4
6	2	1	4	3	2	1	4	3	3	2	1	4
7	1	2	3	4	1	2	3	4	4	1	2	3
8	2	1	4	3	1	3	2	4	2	1	4	3
9	1	2	3	4	2	1	3	4	3	1	2	4
10	1	2	3	4	2	1	3	4	2	1	3	4
11	2	1	3	4	2	1	3	4	1	2	3	4
12	1	2	3	4	1	2	3	4	3	1	2	4
13	1	2	3	4	1	2	4	3	2	1	4	3
14	1	2	3	4	2	3	1	4	3	2	1	4
15	1	2	3	4	1	3	2	4	2	1	4	3
16	1	3	2	4	2	3	1	4	3	2	1	4
17	2	1	3	4	2	1	3	4	1	3	2	4
18	2	1	4	3	3	1	2	4				
19					1	2	3	4				
TOTALS	29	29	61	61	32	34	54	70	38	26	43	63
	$\frac{\Sigma R}{N}$ = 45.00 s = 1024.00 W = 0.63 p = < 0.01				$\frac{\Sigma R}{N}$ = 47.50 s = 971.00 W = 0.54 p = < 0.01				$\frac{\Sigma R}{N}$ = 42.50 s = 713.00 W = 0.49 p = < 0.01			

KEY

A Varied

B Medium

C Long

D Short

(i) Passages considered separately

	A	B	C	D
EARTHQUAKES	29	29	61	61
SOILS	32	34	54	70
TORNADOES	38	26	43	63
TOTALS	99	89	158	194

$\frac{\leq R}{N} = 135.00$

$s = 7422.00$

$W = 0.51$

$p = < 0.01$

KEY

A Varied

B Medium

C Long

D Short

(11) Passages combined

(b) ANALYSIS OF RESIDUALS

FORMULAE

$$\text{Standardised Residual } (e_{ij}) = \frac{(n_{ij} - E_{ij})}{\sqrt{E_{ij}}}$$

$$\text{Variance of Standardised Residual } (V_{ij}) = \left(1 - \frac{n_i}{N}\right)\left(1 - \frac{n_j}{N}\right)$$

$$\text{Adjusted Residual } (d_{ij}) = \frac{e_{ij}}{\sqrt{V_{ij}}}$$

(i) VARIED SENTENCES (A)

Standardised Residuals

	1+2	3	4
EARTHQUAKES	+0.38	+0.48	−1.28
SOILS	+0.59	−0.03	−0.81
TORNADOES	−1.02	−0.46	+2.18

Variance of Standardised Residuals

	1+2	3	4
EARTHQUAKES	0.41	0.38	0.54
SOILS	0.40	0.37	0.53
TORNADOES	0.42	0.39	0.56

(ii) LONG SENTENCES (C)

Standardised Residuals

	1+2	3	4
EARTHQUAKES	−1.79	+0.48	+1.16
SOILS	−0.12	+0.67	−0.69
TORNADOES	+1.97	−1.20	−0.46

Variance of Standardised Residuals

	1+2	3	4
EARTHQUAKES	0.48	0.38	0.47
SOILS	0.47	0.37	0.46
TORNADOES	0.50	0.39	0.48

(c) SOMERS'S d

FORMULA : $dyx = \frac{S}{P + Q + Yo}$ ($S = P - Q$)

P = number of concordant pairs of observations (ie pairs of observations such that their rankings on the two variables are in the same direction.

Q = number of discordant pairs of observations (ie pairs of observations such that their rankings on the two variables are in the reverse direction.

Yo = number of observations tied on the y variable only (ie the dependent variable)

Interpretation

0 = independence

+1 = all data concentrated in upper left to lower right diagonal

(See Everett 1977 : 63-66)

APPENDIX 4

PASSAGE A : SOIL EROSION

PASSAGE B : DEFINITION OF URBAN AREAS

PASSAGE C : THE FOREST COMMUNITY

(a) Each passage with numbered sentences

(b) Each passage with numbered T-Units

(a) NUMBERED SENTENCES

[1]Every geologist is familiar with the erosion cycle.[2]No sooner has an area of land been raised above sea level than it becomes subject to the erosive forces of nature.[3]The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets and then into rivers and out to sea.[4]The frost freezes the rain water in cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust.[5]Blocks of rock dislodged at high levels are brought down by the force of gravity.[6]Alternate heating and cooling of bare rock surfaces causes their disintegration.[7]In the arid regions of the world the wind is a powerful force in removing material from one area to another.[8]All this is natural.[9]But Nature has also provided certain defensive forces.[10]Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks.[11]Slowly but surely different types of soil, with differing profiles, evolve, the main types depending primarily on the climate.[12]The protective soil covering, once it is formed, is held together by the vegetation.[13]Grass and herbaceous plants, with long, branching, tenuous roots, hold firmly together the surface particles.[14]The same is true with the forest cover.
[15]The heaviest tropical downpours, beating on the leaves of the giant trees, reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil.[16]The soil, thus protected by grass, herbs or trees, furnishes a quiet habitat for a myriad varied organisms : earth-worms that importantly modify the soil, bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants.[17]Chemical action is constantly taking place : soil acids attack mineral particles and salts in solution move from one layer in the soil to another.[18]We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions

and the parent rock substances develops.[19] It is a long process, and in many areas there has not been sufficient time, in the geological sense, for completion of the process : the soils are 'immature'.[20] Now let us consider what happens when man the pioneer comes along.[21] He ploughs up the natural grasslands.[22] He removes the various branching roots that have held together the surface particles, so that they are now easily moved by the action of rain and wind.[23] In ploughing, he has mixed together the surface layers, and provided his crops with a medium in which they can grow and develop, and in which natural plant food is present in varying degree, or which can be supplemented by animal or chemical manures.[24] But he has also exposed the soil to the action of the atmosphere.[25] In some parts of the world exposure is useful : in a cold climate the breaking up of the clods of clay by frost action is beneficial.[26] At the other extreme, as we have already pointed out, in tropical climates, by exposure to the atmosphere such rapid chemical action is set up that the natural plant food in the soil is quickly destroyed.[27] Broadly speaking, the position is worse when the natural vegetation is forest.[28] Not only does man, the pioneer, rob the soil of its source of humus, the fallen leaves, but he exposes a soil quite unused to the direct rays of the sun and the direct fall of the rain to the immediate influence of both.[29] For example, much of the upland and west coast of Scotland was once forested with the beautiful Scots fir.[30] The heavy rain, falling on the close pine woods, trickled gradually to the ground and soaked into the soil.[31] Much was evaporated from the leaves, and the floor of the forest, covered with pine leaves, remained comparatively dry, supporting a sparse cover or undergrowth of various shade-loving plants or low shrubs, such as bilberry or heather.[32] When the forests were cut down the heavy rain fell straight on to the surface soil more rapidly than it could drain away.[33] Especially where

there was no steep slope the water was held up, and moisture-loving plants began to flourish, particularly spaghnum, or bog moss.[34] Once the spaghnum was established it acted as a sponge.[35] True, it prevented soil erosion, but it grew and grew until great thicknesses of moss blanketed the whole countryside.[36] Thus, huge stretches of bogland, that known to the botanists as 'blanket bog', extending over wide areas of Scotland and Ireland, are directly due to man's action.

From Stamp, L. D. (1960) Our Developing World Faber & Faber 99-101.

(a) NUMBERED SENTENCES

[1]Let us consider the various definitions of urban settlements.[2] The definitions used in legal and administrative documents will tell us precisely what we mean by Melbourne, Australia or Topeka, Kansas. [3]Unfortunately, the legal and administrative borders of cities are often a historical or constitutional legacy.[4] Typically, the legal city has fixed boundaries that survive long after urban development have exceeded these bounds.[5] Thus the legal city is often 'under-bounded'. [6]Parts of the urban area may remain outside the city, but share a common boundary with it.[7] Beverley Hills, completely surrounded by the city of Los Angeles, is a case in point.[8] In England, where administrative regions continue to have a strong administrative quality, some boroughs still have a municipal status that is a legacy of their former importance, and is out of line with their present small size.[9] The number of inhabitants an area must have to be considered urban also varies from country to country.[10] In Iceland places with a few hundred people are termed urban, whereas in the Netherlands a population of 20,000 is needed.[11] A second approach to defining urban settlements is to ignore the legal boundaries and try to define each settlement in terms of its physical structure.[12] For example, we might define a settlement on the basis of a continuous distribution of housing, or population above a certain density, or the intensity of traffic.[13] But there are difficulties here too.[14] What do we mean by 'continuous housing', and what happens when different definitions don't all give the same answer?[15] Figure 14.2 presents some different definitions of New York, based on both its legal boundaries and its physical structure.[16] Note that New York City itself is only a small part of the continuous urban sprawl that is Greater New York.[17] The mismatch between the legal and the physical city becomes vitally important when the legal city, with its static or declining population and limited tax base, has to provide public services, like

transport or police, for the millions of commuters who cross its boundaries for work each day.[18] As the discrepancy between the legal and economic boundaries of the city becomes worse, the pressure for some form of revenue-sharing or boundary adjustment grows.[19] This discrepancy also affects our ability to answer even the simplest questions about the size of the city.[20] To take an extreme case, the 'legal city' of Sydney, Australia, in 1955 had a population of only 193,000, while the 'built-up' area of Sydney had a population of 1,869,000.[21] This difference of over nine times in size is unusual, but important enough to make the definition of settlement a matter of concern.[22] As a result of this problem, international and indeed intranational definitions of urban settlements are being standardized.[23] One definition of world metropolitan areas by demographer KingsleyDavis runs to twelve pages, including two pages on difficult cases.[24] In the United States, the concept of a Standard Metropolitan Statistical Area (SMSA) was introduced in 1960, so that metropolitan areas could be defined realistically by using three criteria.[25] First, a population criterion : each SMSA must include one central city with 50,000 or more inhabitants.[26] Special rules allow contiguous cities (ie those directly adjoining each other) and nearby cities (within twenty miles of each other) to be combined.[27] Second, the metropolitan character of an area is taken into account.[28] At least 75 percent of the labour force of the county must be employed by non-agricultural industries.[29] Other criteria for SMSAs relate to population density, the contiguity of townships and ratios between the non-agricultural labour forces making up the unit.[30] Finally, the integration of the areas that constitute the SMSA is considered.[31] Counties are integrated within the county containing a central city if 15 percent of the workers in the county live in the city.[32] This measure of integration can be supplemented by other measures based on the market area, newspaper subscriptions, retail trade, public transport, and the like.[33] Despite their apparent compre-

hensiveness, the SMSA definitions have still not solved the problem of urban boundaries.[34] An improved definition using county blocks and commuting data has been suggested by a team of Chicago geographers.[35] Their recommendations were threefold.[36] First, counties of equivalent units were to be retained as the basic building blocks of any system.[37] Second, counties were to be classified into functional economic areas (FEAs) on the basis of county to county commuting data.[38] An FEA would consist of all the counties in which the proportion of resident workers who commuted to a given central county (usually containing a city of 50,000 or more inhabitants) exceeded the proportion who commuted to alternative central counties.[39] Third, FEAs were to be grouped into consolidated urban regions (CURs) when two or more FEAs sent at least 5 percent of the workers in the central county of one commuting area to the central county of others.

From Haggett, P. (1975) Geography : A Modern Synthesis Harper & Row (International Edition) 352-355

(a) NUMBERED SENTENCES

[1]In the forest community, every individual strives to obtain for itself and its offspring the space and the nourishment necessary for life. [2]This involves competition.[3]The struggle for existence, which plays such an important part in nature, is exemplified in a most striking manner in the forest.[4]Above ground it resolves itself mainly into a struggle for light.[5]Every tree competes with its neighbours for as large a share as possible for the light falling upon the forest from the sky, by striving to thrust its crown, with its assimilating organs the leaves, above those of its neighbours.[6]The individuals which fail to hold their own in the struggle are doomed to destruction sooner or later.[7]Their crowns are overshaded, their assimilation drops, and, left behind in the semi-darkness below, they ultimately fail to nourish themselves sufficiently to maintain life.[8]Owing to the widespread roots of forest trees, competition also goes on in the soil for water and soil nutrients.[9]The smaller plants, bushes, herbs, mosses, etc, are also involved in competition.[10]While they have, as a rule, to submit to the dominence of the trees, they compete amongst themselves and with the younger progeny of the trees for such light as can reach them, and for their share of the moisture and plant food in the soil.[11]In a true community, however, besides competition there is also cooperation among the members.[12]The trees protect and support each other against the wind, and their crowns unite to form a canopy which shades the soil and so keeps injurious weeds out of the forest.[13]They protect, by the shelter they give, not only their own but their neighbours' offspring from destructive frosts.[14]They produce within the forest a climate quite different from that outside.[15]They also produce distinctive soil conditions, which are almost essential for many members of the community, and are, for others at any rate, very beneficial.[16]The long continued interaction of competition and cooperation tends to produce ultimately a condition of equilibrium in the forest community, which gives it

stability and permanence.[17] When this condition is reached, the forest is said to be the climax type of forest for that particular locality.[18] The species and their relations to one another best adapted to the conditions of the locality have been established.[19] The climax type is the end of a succession of types of vegetation which may have occupied the site in the past, each of which, by modifying the climate and soil of the site, has paved the way for its supersession by another type.[20] New species have found conditions favourable for establishing themselves, and more or less completely ousted the old.[21] The climax type depends on the climate and soil, and the particular species of plants and animals which have had the opportunity of entering the community, and their actions and reactions on each other.[22] The equilibrium established in the forest community is not static but dynamic.[23] It constantly undergoes fluctuations through variations in the weather from year to year, which favour or discourage certain members or sets of members.[24] There are, however, always compensating factors.[25] Wet years are succeeded by dry years, warm seasons by cold ones.[26] If one set of members multiplies unduly, its enemies also increase, and the over-produced species is brought back to normal.[27] The mechanism of the biotic equilibrium is so constructed that it always re-establishes itself, provided there are no over-powerful or persistent interferences from outside.[28] The most important of these interferences are those for which man is responsible.[29] He has introduced other members into the forest community, such as new species of trees, grazing animals, etc, and, with axe and fire, has introduced new factors into the environment, favouring some, and injuring or destroying other sets of members, with varying effects on the forest.[30] Some of these effects are direct and evident, some indirect, and not immediately obvious, but nevertheless of far-reaching consequences.[31] It is therefore necessary that silviculture should be based, not only on a knowledge of the natural relationships of the forest to its environment, but also of the responses of the forest to human actions of various kinds.

From Thomson, T. & M. R. K. Jerram (1948) <u>An Outline of Forestry</u>
Allen & Unwin Ltd

(b) NUMBERED T-UNITS

[1]Every geologist is familiar with the erosion cycle.|[2]No sooner has an area of land been raised above sea level than it becomes subject to the erosive forces of nature.|[3]The rain beats down on the ground|[4]and washes away the finer particles, sweeping them into rivulets and then into rivers and out to sea.|[5]The frost freezes the rain water in cracks in the rocks|[6]and breaks up even the hardest of the constituents of the earth's crust.|[7]Blocks of rock dislodged at high levels are brought down by the force of gravity.|[8]Alternate heating and cooling of bare rock surfaces causes their disintegration.|[9]In the arid regions of the world the wind is a powerful force in removing material from one area to another.|[10]All this is natural.|[11]But Nature has also provided certain defensive forces.|[12]Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks.|[13]Slowly but surely different types of soil, with differing profiles, evolve, the main types depending primarily on the climate.|[14]The protective soil covering, once it is formed, is held together by the vegetation.|[15]Grass and herbaceous plants, with long, branching, tenuous roots, hold firmly together the surface particles.|[16]The same is true with the forest cover.|[17]The heaviest tropical downpours, beating on the leaves of the giant trees, reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil.|[18]The soil, thus protected by grass, herbs or trees, furnishes a quiet habitat for a myriad varied organisms :|[19]earth-worms that importantly modify the soil,|[20]bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants.|[21]Chemical action is constantly taking place :|[22]soil acids attack mineral particles|[23]and salts in solution move from one layer in the soil to another.|[24]We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions

and the parent rock substances develops.|25 It is a long process,|26 and in many areas there has not been sufficient time, in the geological sense, for completion of the process :|27 the soils are 'immature'.|28 Now let us consider what happens when man the pioneer comes along.|29 He ploughs up the natural grasslands.|30 He removes the various branching roots that have held together the surface particles, so that they are now easily moved by the action of rain and wind.|31 In ploughing, he has mixed together the surface layers,|32 and provided his crops with a medium in which they can grow and develop, and in which natural plant food is present in varying degree, or which can be supplemented by animal or chemical manures.|33 But he has also exposed the soil to the action of the atmosphere.|34 In some parts of the world exposure is useful :|35 in a cold climate the breaking up of the clods of clay by frost action is beneficial.|36 At the other extreme, as we have already pointed out, in tropical climates, by exposure to the atmosphere such rapid chemical action is set up that the natural plant food in the soil is quickly destroyed.|37 Broadly speaking, the position is worse when the natural vegetation is forest.|38 Not only does man, the pioneer, rob the soil of its source of humus, the fallen leaves,|39 but he exposes a soil quite unused to the direct rays of the sun and the direct fall of the rain to the immediate influence of both.|40 For example, much of the upland and west coast of Scotland was once forested with the beautiful Scots fir.|41 The heavy rain, falling on the close pine woods, trickled gradually to the ground|42 and soaked into the soil.|43 Much was evaporated from the leaves,|44 and the floor of the forest, covered with pine leaves, remained comparatively dry, supporting a sparse cover or undergrowth of various shade-loving plants or low shrubs, such as bilberry or heather.|45 When the forests were cut down the heavy rain fell straight on to the surface soil more rapidly than it could drain away.|46 Especially where

there was no steep slope the water was held up,|47 and moisture-loving plants began to flourish, particularly spaghnum, or bog moss.|48 Once the spaghnum was established it acted as a sponge.|49 True, it prevented soil erosion,|50 but it grew and grew until great thicknesses of moss blanketed the whole countryside.|51 Thus, huge stretches of bogland, that known to the botanists as 'blanket bog', extending over wide areas of Scotland and Ireland, are directly due to man's action.|

From Stamp, L. D. (1960) Our Developing World Faber & Faber 99-101.

(b) NUMBERED T-UNITS

[1]Let us consider the various definitions of urban settlements.|[2]The definitions used in legal and administrative documents will tell us precisely what we mean by Melbourne, Australia or Topeka, Kansas.|[3]Unfortunately, the legal and administrative borders of cities are often a historical or constitutional legacy.|[4]Typically, the legal city has fixed boundaries that survive long after urban development have exceeded these bounds.|[5]Thus the legal city is often 'under-bounded'.|[6]Parts of the urban area may remain outside the city,|[7]but share a common boundary with it.|[8]Beverley Hills, completely surrounded by the city of Los Angeles, is a case in point.|[9]In England, where administrative regions continue to have a strong administrative quality, some boroughs still have a municipal status that is a legacy of their former importance, and is out of line with their present small size.|[10]The number of inhabitants an area must have to be considered urban also varies from country to country.|[11]In Iceland places with a few hundred people are termed urban,|[12]whereas in the Netherlands a population of 20,000 is needed.|[13]A second approach to defining urban settlements is to ignore the legal boundaries and try to define each settlement in terms of its physical structure.|[14]For example, we might define a settlement on the basis of a continuous distribution of housing, or population above a certain density, or the intensity of traffic.|[15]But there are difficulties here too.|[16]What do we mean by 'continuous housing',|[17]and what happens when different definitions don't all give the same answer?|[18]Figure 14.2 presents some different definitions of New York, based on both its legal boundaries and its physical structure.|[19]Note that New York City itself is only a small part of the continuous urban sprawl that is Greater New York.|[20]The mismatch between the legal and the physical city becomes vitally important when the legal city, with its static or declining population and limited tax base, has to provide public services, like

transport or police, for the millions of commuters who cross its boundaries for work each day.|21 As the discrepancy between the legal and economic boundaries of the city becomes worse, the pressure for some form of revenue-sharing or boundary adjustment grows.|22 This discrepancy also affects our ability to answer even the simplest questions about the size of the city.|23 To take an extreme case, the 'legal city' of Sydney, Australia, in 1955 had a population of only 193,000,|24 while the 'built-up' area of Sydney had a population of 1,869,000.|25 This difference of over nine times in size is unusual,|26 but important enough to make the definition of settlement a matter of concern.|27 As a result of this problem, international and indeed intranational definitions of urban settlements are being standardized.|28 One definition of world metropolitan areas by demographer KingsleyDavis runs to twelve pages, including two pages on difficult cases.|29 In the United States, the concept of a Standard Metropolitan Statistical Area (SMSA) was introduced in 1960, so that metropolitan areas could be defined realistically by using three criteria.|30 First, a population criterion :|31 each SMSA must include one central city with 50,000 or more inhabitants.|32 Special rules allow contiguous cities (ie those directly adjoining each other) and nearby cities (within twenty miles of each other) to be combined.|33 Second, the metropolitan character of an area is taken into account.|34 At least 75 percent of the labour force of the county must be employed by non-agricultural industries.|35 Other criteria for SMSAs relate to population density, the contiguity of townships and ratios between the non-agricultural labour forces making up the unit.|36 Finally, the integration of the areas that constitute the SMSA is considered.|37 Counties are integrated within the county containing a central city if 15 percent of the workers in the county live in the city.|38 This measure of integration can be supplemented by other measures based on the market area, newspaper subscriptions, retail trade, public transport, and the like.|39 Despite their apparent compre-

hensiveness, the SMSA definitions have still not solved the problem of urban boundaries.[40] An improved definition using county blocks and commuting data has been suggested by a team of Chicago geographers.[41] Their recommendations were threefold.[42] First, counties of equivalent units were to be retained as the basic building blocks of any system.[43] Second, counties were to be classified into functional economic areas (FEAs) on the basis of county to county commuting data.[44] An FEA would consist of all the counties in which the proportion of resident workers who commuted to a given central county (usually containing a city of 50,000 or more inhabitants) exceeded the proportion who commuted to alternative central counties.[45] Third, FEAs were to be grouped into consolidated urban regions (CURs) when two or more FEAs sent at least 5 percent of the workers in the central county of one commuting area to the central county of others.

From : Haggett, P. (1975) Geography : A Modern Synthesis Harper & Row (International Edition) 352-355.

(b) NUMBERED T-UNITS

[1]In the forest community, every individual strives to obtain for itself and its offspring the space and the nourishment necessary for life.| [2]This involves competition.| [3]The struggle for existence, which plays such an important part in nature, is exemplified in a most striking manner in the forest.| [4]Above ground it resolves itself mainly into a struggle for light.| [5]Every tree competes with its neighbours for as large a share as possible for the light falling upon the forest from the sky, by striving to thrust its crown, with its assimilating organs the leaves, above those of its neighbours.| [6]The individuals which fail to hold their own in the struggle are doomed to destruction sooner or later.| [7]Their crowns are overshaded,| [8]their assimilation drops,| [9]and, left behind in the semi-darkness below, they ultimately fail to nourish themselves sufficiently to maintain life.| [10]Owing to the widespread roots of forest trees, competition also goes on in the soil for water and soil nutrients.| [11]The smaller plants, bushes, herbs, mosses, etc, are also involved in competition.| [12]While they have, as a rule, to submit to the dominence of the trees, they compete amongst themselves and with the younger progeny of the trees for such light as can reach them, and for their share of the moisture and plant food in the soil.| [13]In a true community, however, besides competition there is also cooperation among the members.| [14]The trees protect and support each other against the wind,| [15]and their crowns unite to form a canopy which shades the soil and so keeps injurious weeds out of the forest.| [16]They protect, by the shelter they give, not only their own but their neighbours' offspring from destructive frosts.| [17]They produce within the forest a climate quite different from that outside.| [18]They also produce distinctive soil conditions, which are almost essential for many members of the community, and are, for others at any rate, very beneficial.| [19]The long continued interaction of competition and cooperation tends to produce ultimately a condition of equilibrium in the forest community, which gives it

stability and permanence.|20 When this condition is reached, the forest is said to be the climax type of forest for that particular locality.|21 The species and their relations to one another best adapted to the conditions of the locality have been established.|22 The climax type is the end of a succession of types of vegetation which may have occupied the site in the past, each of which, by modifying the climate and soil of the site, has paved the way for its supersession by another type.|23 New species have found conditions favourable for establishing themselves,|24 and more or less completely ousted the old.|25 The climax type depends on the climate and soil, and the particular species of plants and animals which have had the opportunity of entering the community, and their actions and reactions on each other.|26 The equilibrium established in the forest community is not static but dynamic.|27 It constantly undergoes fluctuations through variations in the weather from year to year, which favour or discourage certain members or sets of members.|28 There are, however, always compensating factors.|29 Wet years are succeeded by dry years,|30 warm seasons by cold ones.|31 If one set of members multiplies unduly, its enemies also increase,|32 and the over-produced species is brought back to normal.|33 The mechanism of the biotic equilibrium is so constructed that it always re-establishes itself, provided there are no over-powerful or persistent interferences from outside.|34 The most important of these interferences are those for which man is responsible.|35 He has introduced other members into the forest community, such as new species of trees, grazing animals, etc,|36 and, with axe and fire, has introduced new factors into the environment, favouring some, and injuring or destroying other sets of members, with varying effects on the forest.|37 Some of these effects are direct and evident,|38 some indirect, and not immediately obvious, but nevertheless of far-reaching consequences.|39 It is therefore necessary that silviculture should be based, not only on a knowledge of the natural relationships of the forest to its environment, but also of the responses of the forest to human actions of various kinds.

From Thomson, T & M. R. K. Jerram (1948) An Outline of Forestry Allen & Unwin Ltd.

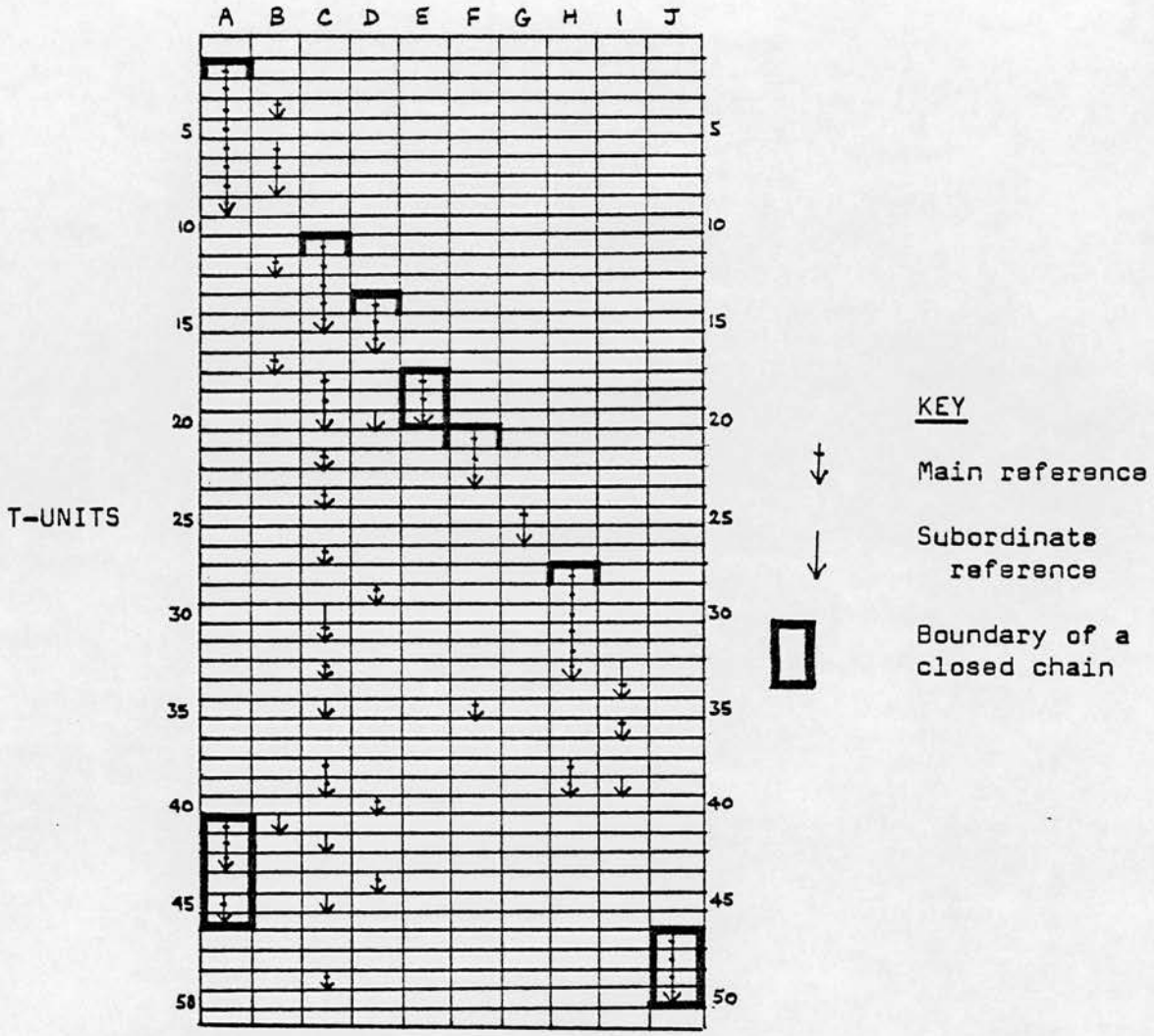
APPENDIX 5

(a) LEXICAL COHESION CHAINS

(b) MOVEMENT FROM PARTICULAR TO GENERAL

(a) LEXICAL COHESION CHAINS

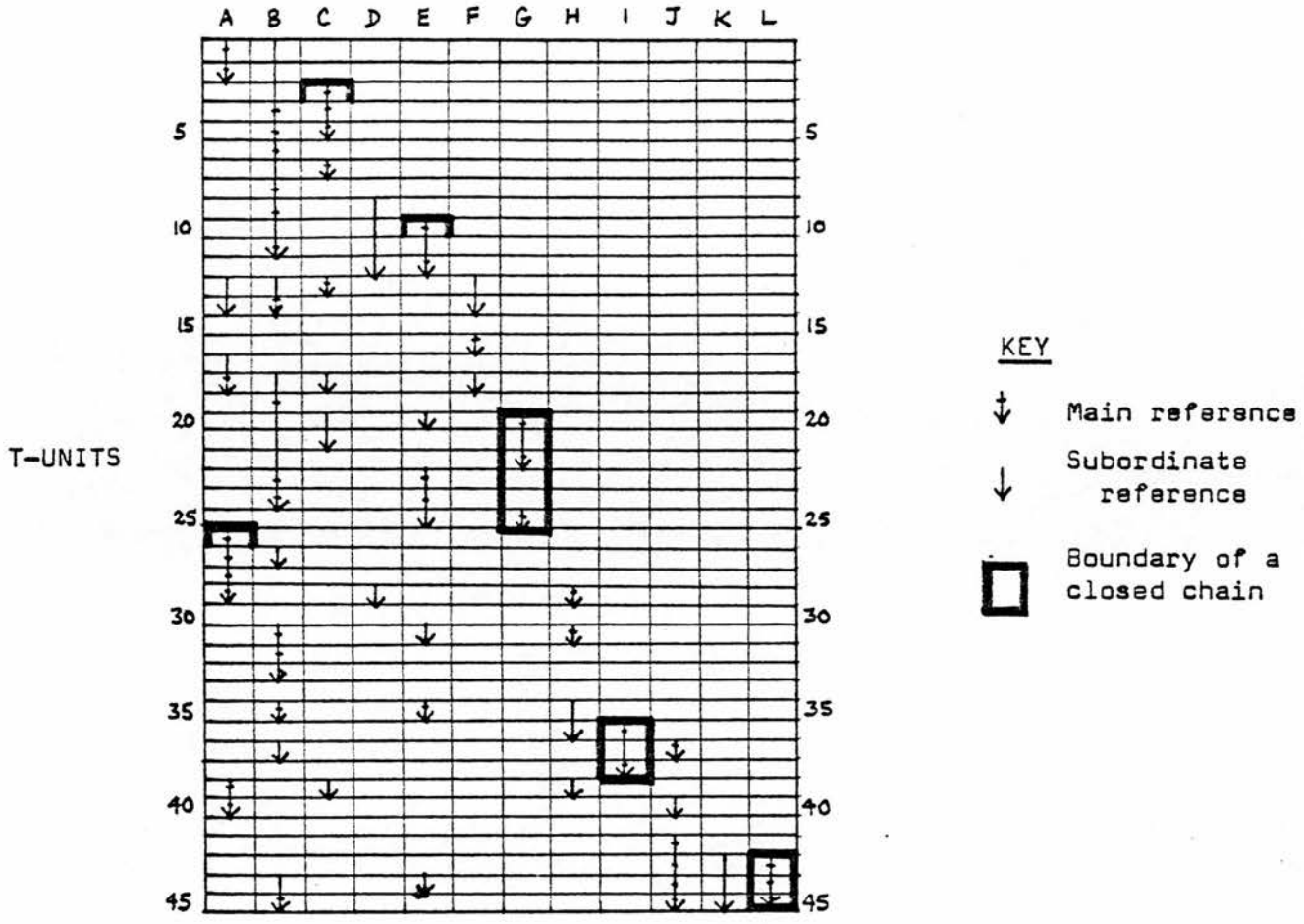
APPENDIX 5



(i) Passage A : Soil Erosion

LEXICAL COHESION CHAINS

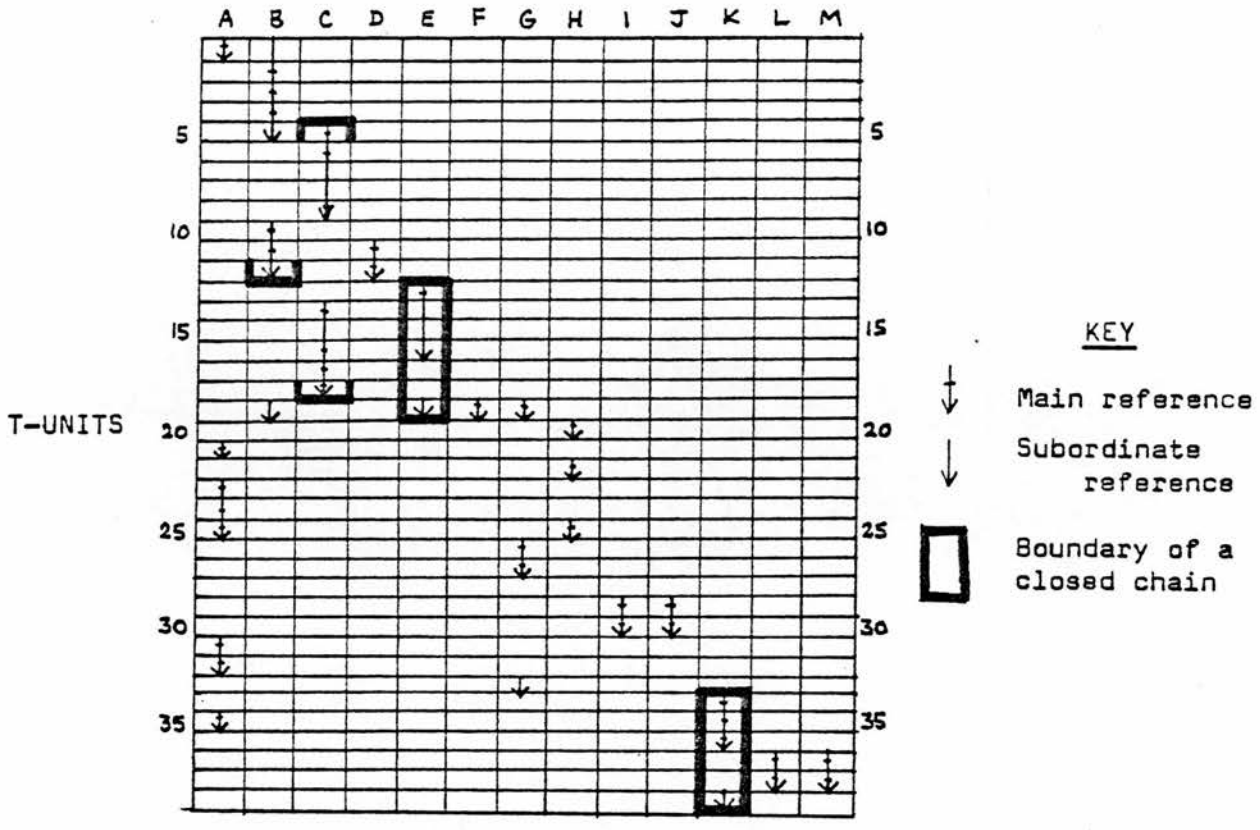
A Erosive forces
B Surface particles
C Defensive forces : soil
D Defensive forces : vegetation
E Organisms in the soil
F Chemical processes
G A long process
H Man the pioneer
I Exposure
J Moisture-loving plants



(ii) Passage B : Definition of Urban Areas

LEXICAL COHESION CHAINS

A The various definitions of urban settlements
B Urban settlements
C Legal and administrative borders of cities
D In England (followed by country references)
E The number of inhabitants
F Physical structure
G The mismatch between the legal and physical city
H The concept of an SMSA
I The integration of the areas that constitute the SMSAs
J Counties
K Commuting data
L FEAs

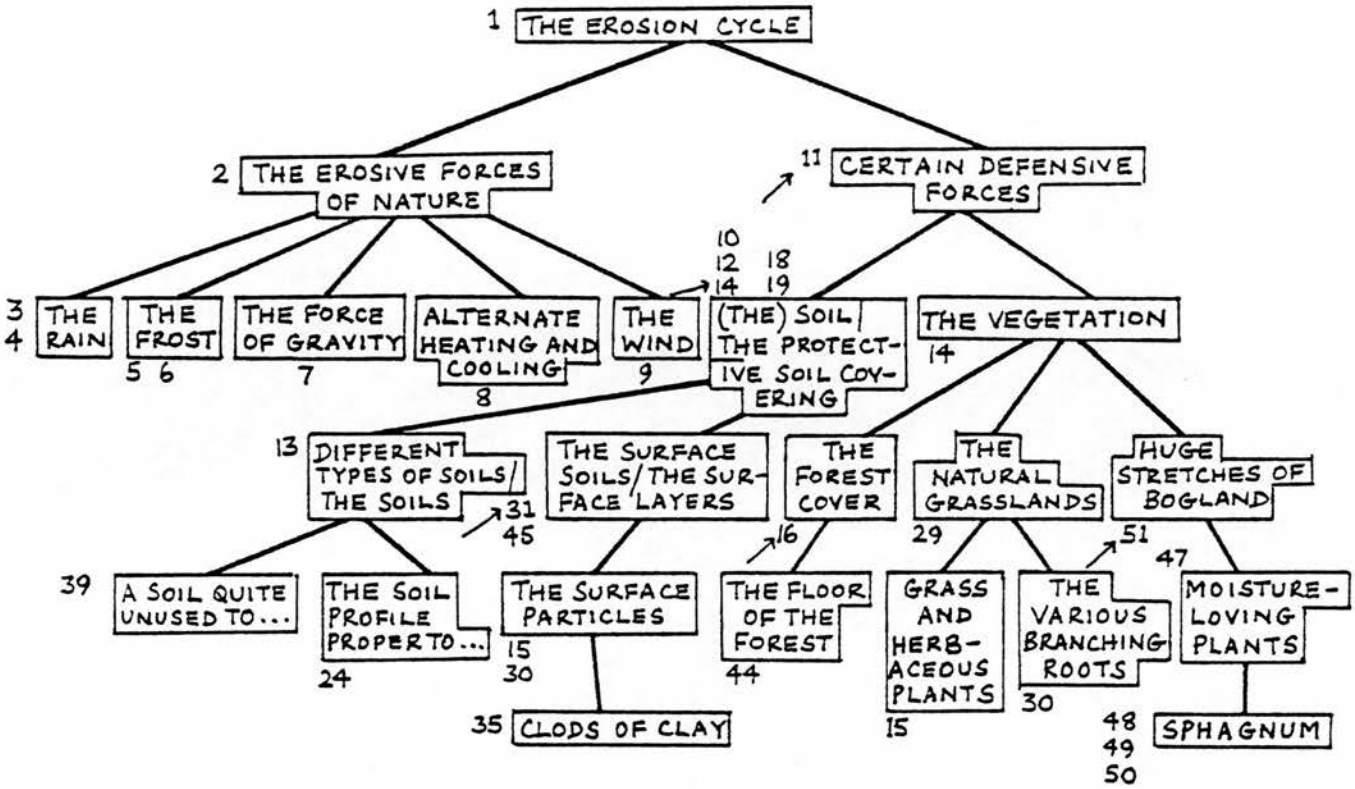


(iii) Passage C : The Forest Community

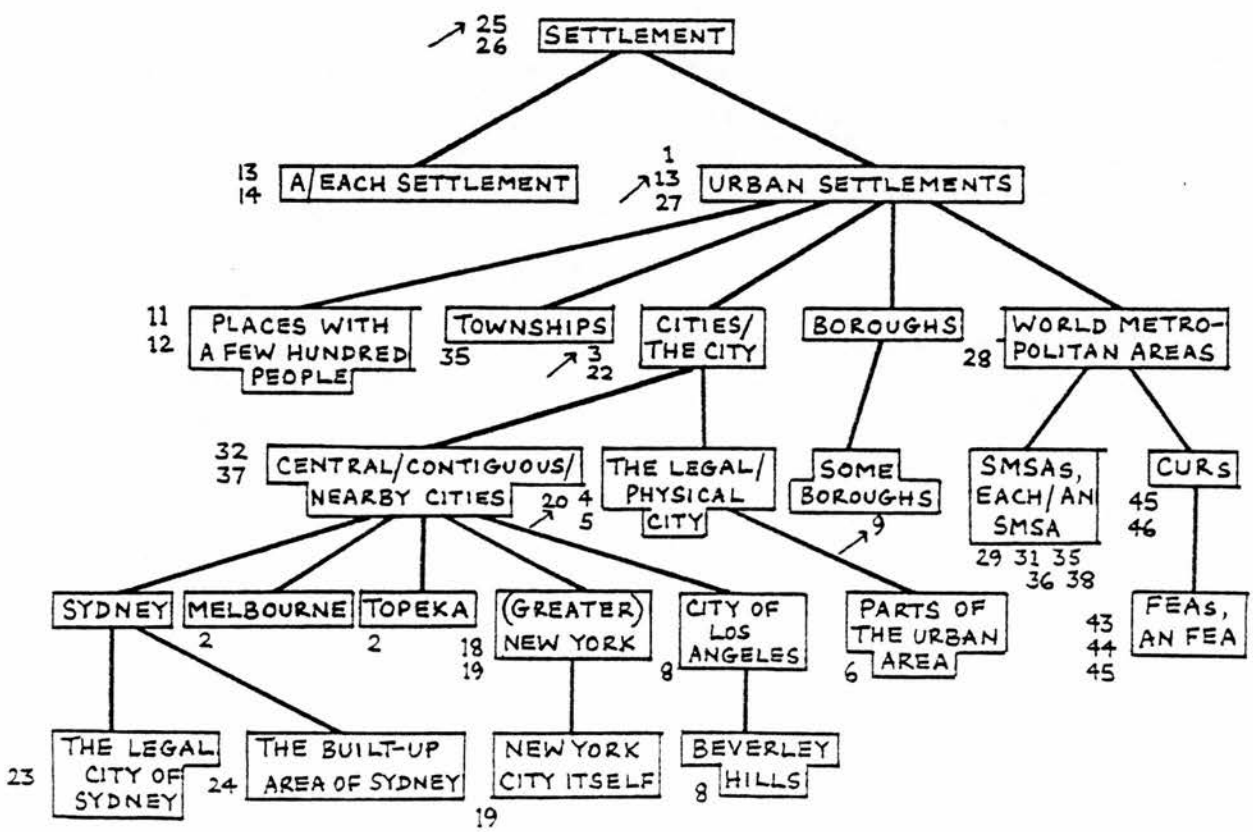
LEXICAL COHESION CHAINS

A Individuals and species
B Competition
C The trees
D The smaller plants
E Co-operation
F The interaction of competition and co-operation
G The condition of equilibrium
H The climax type of forest
I Climatic conditions
J Years and seasons
K Man
L The effects of man
M Direct and indirect

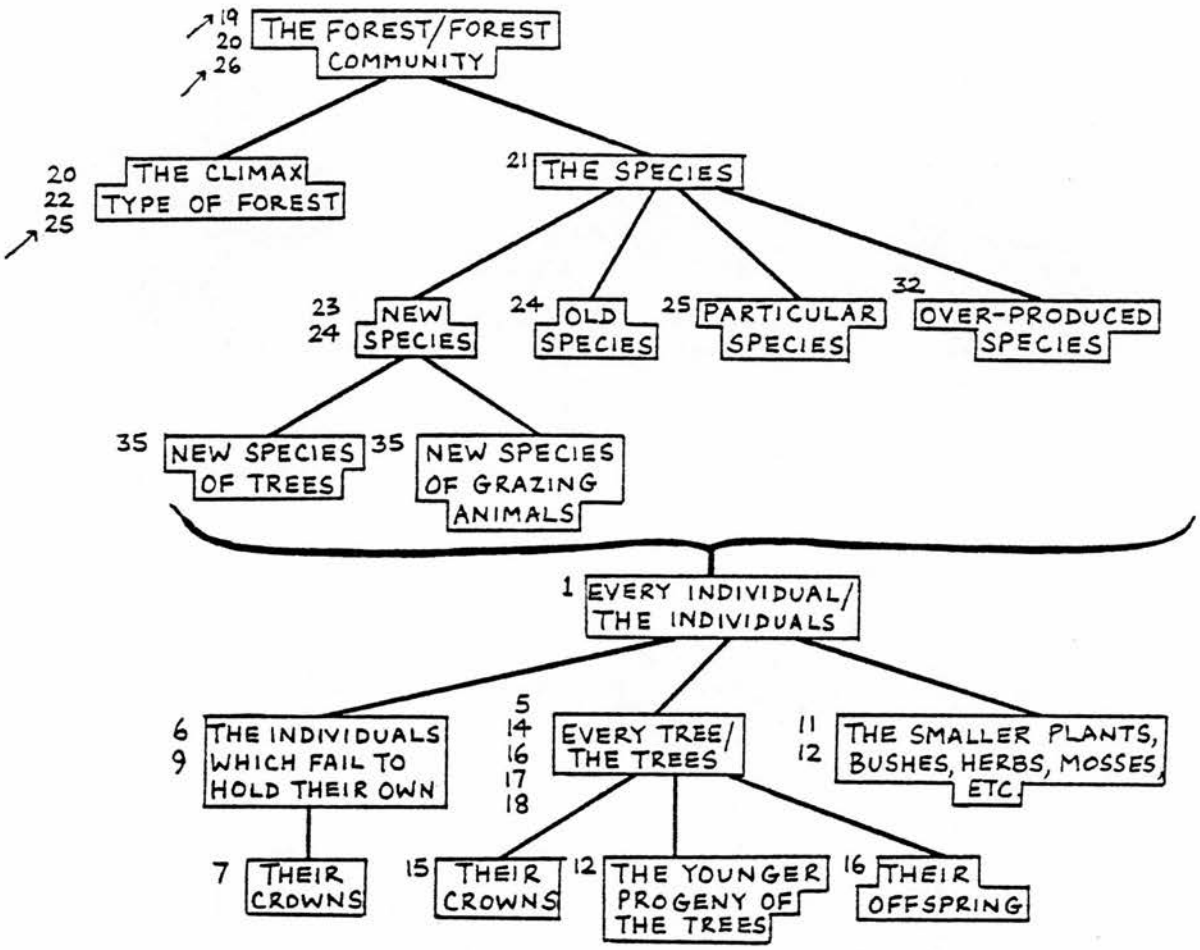
(b) MOVEMENT FROM PARTICULAR TO GENERAL



(i) Passage A : Soil Erosion



(ii) Passage B : Definition of Urban Areas



(iii) Passage C : The Forest Community

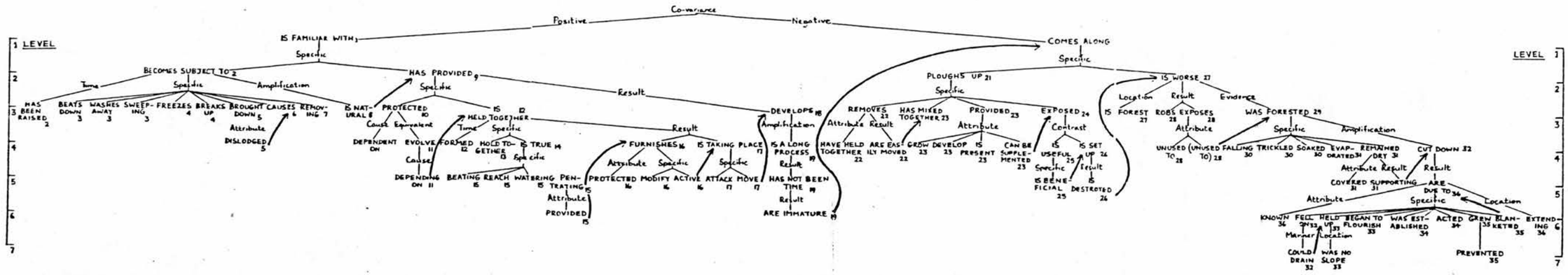
APPENDIX 6

AN ANALYSIS OF THE THREE PASSAGES ACCORDING TO A MODIFIED VERSION OF GRIMES' SEMANTIC GRAMMAR OF PROPOSITIONS

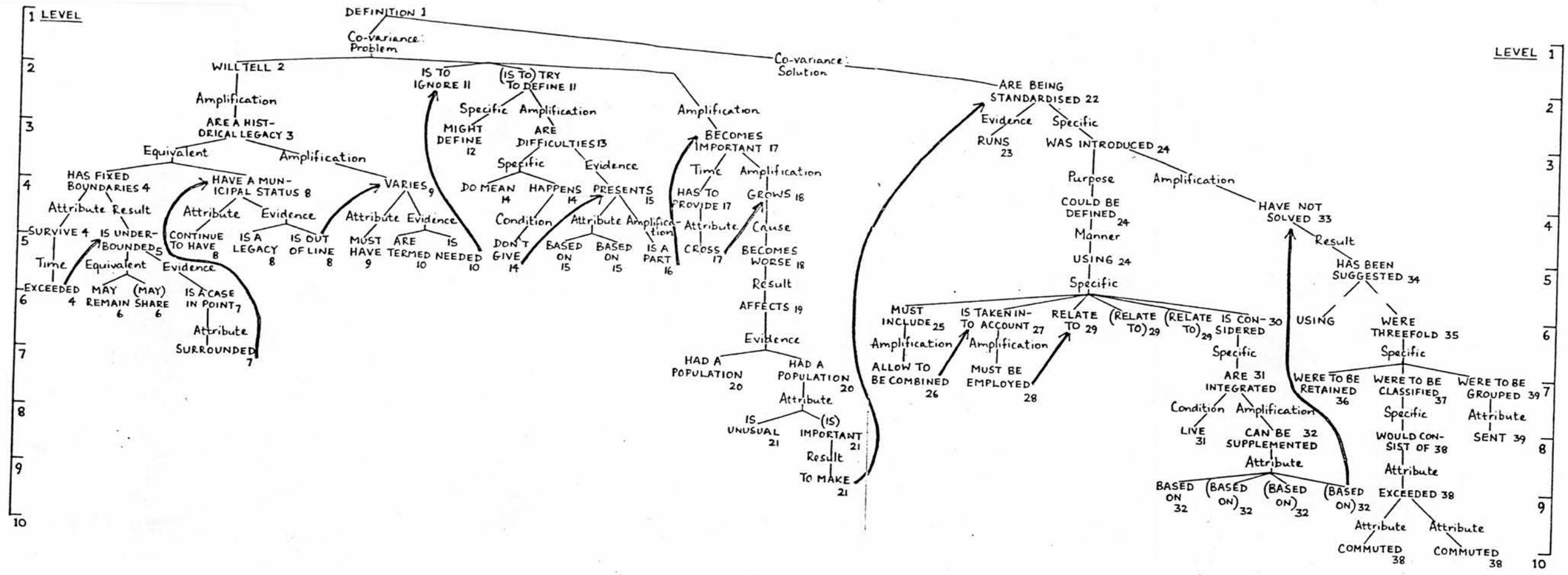
(a) Passage A : Soil Erosion

(b) Passage B : Definition of Urban Areas

(c) Passage C : The Forest Community

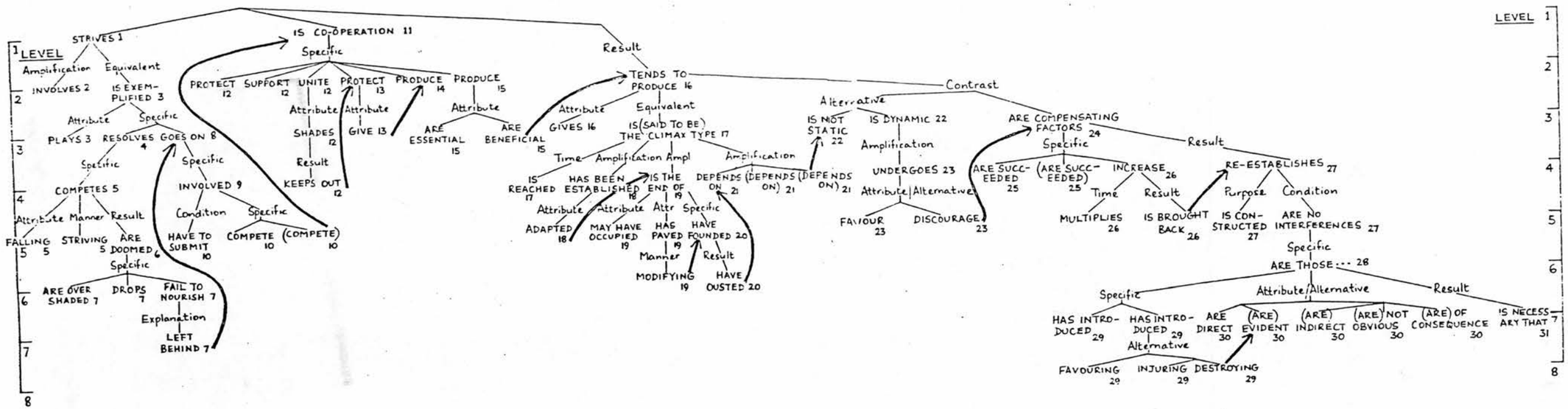


PASSAGE A (SOIL EROSION) : Hierarchical tree diagram of the content structure



PASSAGE B (DEFINITION OF URBAN AREAS) : Hierarchical tree diagram of the content structure

APPENDIX 6c



PASSAGE C (THE FOREST COMMUNITY) : Hierarchical tree diagram of the content structure

APPENDIX 7

EXPERIMENT 3

(a) Instructions given to the subjects

(b) Raw results

(a) INSTRUCTIONS

PASSAGE A

1. Divide the passage into <u>two</u> paragraphs. Mark your division as follows : ||
2. Divide each of these paragraphs into <u>two</u> further paragraphs. Mark your divisions as follows : | – (You now have a total of four paragraphs)
3. Add either <u>two</u> or <u>three</u> extra paragraphs. Mark your divisions as follows : ¦ (You now have a total of either six or seven paragraphs)
4. Although you may not wish to make further <u>paragraph</u> divisions, it is still possible to identify smaller segments of the passage. Further divide the passage so that you have a final total of <u>up to sixteen</u> divisions (including, of course, the boundaries you have already identified) Mark these final divisions as follows : ⋮

PASSAGE B

1. Divide the passage into <u>two</u> paragraphs. Mark your division as follows : ||
2. Divide each of these paragraphs into <u>three</u> further paragraphs. Mark your divisions as follows : | (You now have a total of six paragraphs)
3. Although you may not wish to make further <u>paragraph</u> divisions, it is still possible to identify smaller segments of the passage. Further divide the passage so that you have a final total of <u>up to twenty</u> divisions (including the boundaries you have already identified) Mark these final divisions as follows : ¦.

PASSAGE C

1. Divide the passage into <u>two</u> paragraphs. Mark your division as follows : ||
2. Divide each of these paragraphs into <u>two</u> further paragraphs. Mark your divisions as follows : | (You now have a total of four paragraphs)
3. Add either <u>two</u> or <u>three</u> extra paragraphs. Mark your divisions as follows : ¦ (You now have a total of six or seven paragraphs)
4. Although you may not wish to make further <u>paragraph</u> divisions, it is still possible to identify smaller segments of the passage. Further divide the passage so that you have a total of <u>up to fourteen</u> divisions (including the boundaries you have already identified) Mark these final divisions as follows : ⋮

(b) EXPERIMENT 3 : RAW RESULTS

(i) FIRST DIVISION

PASSAGE A

1	-	11	-	21	-	31	-
2	-	12	-	22	-	32	-
3	-	13	-	23	-	33	-
4	-	14	-	24	-	34	-
5	-	15	-	25	-	35	-
6	-	16	-	26	-	36	-
7	-	17	-	27	-		
8	-	18	-	28	-		
9	-	19	-	29	-		
10	-	20	19 (100%)	30	-		

PASSAGE B

1	-	11	2 (11.11%)	21	-	31	-
2	-	12	-	22	13 (72.22%)	32	-
3	-	13	-	23	-	33	-
4	-	14	-	24	-	34	-
5	-	15	-	25	-	35	-
6	-	16	-	26	-	36	-
7	-	17	2 (11.11%)	27	-	37	-
8	-	18	-	28	-	38	-
9	-	19	1 (1.56%)	29	-	39	-
10	-	20	-	30	-		

PASSAGE C

1	-	11	1 (5.56%)	21	-	31	-
2	-	12	-	22	2 (11.11%)		
3	-	13	-	23	-		
4	-	14	-	24	-		
5	-	15	-	25	-		
6	-	16	10 (55.56%)	26	-		
7	-	17	1 (5.56%)	27	1 (5.56%)		
8	-	18	1 (5.56%)	28	1 (5.56%)		
9	-	19	1 (5.56%)	29	-		
10	-	20	-	30	-		

(ii) SECOND DIVISION

PASSAGE A

1	-	11	-	21	-	31	-
2	-	12	1 (5.26%)	22	-	32	-
3	-	13	-	23	-	33	-
4	-	14	2 (10.53%)	24	4 (21.05%)	34	-
5	-	15	-	25	1 (5.26%)	35	-
6	-	16	1 (5.26%)	26	-	36	-
7	-	17	1 (5.26%)	27	13 (68.42%)		
8	5 (26.32%)	18	1 (5.26%)	28	1 (5.26%)		
9	7 (36.84%)	19	-	29	-		
10	1 (5.26%)	20	-	30	-		

PASSAGE B

1	-	11	15 (83.33%)	21	-	31	-
2	-	12	-	22	2 (11.11%)	32	-
3	2 (11.11%)	13	1 (5.56%)	23	-	33	6 (33.33%)
4	1 (5.56%)	14	-	24	10 (55.56%)	34	10 (55.56%)
5	1 (5.56%)	15	1 (5.56%)	25	4 (22.22%)	35	-
6	-	16	-	26	-	36	-
7	-	17	10 (55.56%)	27	-	37	-
8	3 (16.67%)	18	2 (11.11%)	28	-	38	-
9	2 (11.11%)	19	1 (5.56%)	29	-	39	-
10	-	20	-	30	1 (5.56%)		

PASSAGE C

1	-	11	14 (77.78%)	21	-	31	3 (16.67%)
2	-	12	-	22	10 (55.56%)		
3	-	13	-	23	-		
4	-	14	-	24	1 (5.56%)		
5	-	15	-	25	-		
6	-	16	1 (5.56%)	26	-		
7	-	17	1 (5.56%)	27	2 (11.11%)		
8	1 (5.56%)	18	-	28	2 (11.11%)		
9	-	19	1 (5.56%)	29	-		
10	-	20	-	30	-		

(iii) THIRD DIVISION

PASSAGE A

1	-	11	-	21	-	31	-
2	-	12	3 (15.79%)	22	-	32	5 (26.32%)
3	1 (5.26%)	13	-	23	-	33	-
4	-	14	2 (10.53%)	24	5 (26.32%)	34	-
5	1 (5.26%)	15	-	25	-	35	-
6	-	16	6 (31.58%)	26	-	36	1 (5.26%)
7	1 (5.26%)	17	2 (10.53%)	27	5 (26.32%)		
8	1 (5.26%)	18	8 (42.11%)	28	-		
9	1 (5.26%)	19	-	29	2 (10.53%)		
10	-	20	-	30	1 (5.26%)		

PASSAGE C

1	-	11	3 (16.67%)	21	1 (5.56%)	31	3 (16.67%)
2	-	12	-	22	2 (11.11%)		
3	-	13	-	23	-		
4	3 (16.67%)	14	-	24	-		
5	-	15	-	25	-		
6	1 (5.56%)	16	3 (16.67%)	26	-		
7	-	17	-	27	5 (27.78%)		
8	6 (33.33%)	18	-	28	6 (33.33%) *		
9	1 (5.56%)	19	1 (5.56%)	29	-		
10	-	20	-	30	1 (5.56%)		

* added after weighted score

(iv) FOURTH DIVISION

PASSAGE A

1	-	11	1 (5.26%)	21	1 (5.26%)	31	1 (5.26%)
2	1 (5.26%)	12	11 (57.89%)	22	-	32	9 (47.37%)
3	7 (36.84%)	13	-	23	5 (26.32%)	33	-
4	6 (31.58%)	14	9 (47.37%)	24	6 (31.58%)	34	1 (5.26%)
5	3 (15.79%)	15	-	25	4 (21.05%)	35	2 (10.53%)
6	3 (15.79%)	16	9 (47.37%)	26	4 (21.05%)	36	7 (36.84%)
7	6 (31.58%)	17	12 (63.16%)	27	-		
8	1 (5.26%)	18	7 (36.84%)	28	-		
9	3 (15.79%)	19	-	29	12 (63.16%)		
10	2 (10.53%)	20	-	30	-		

PASSAGE B

1	-	11	1 (5.56%)	21	-	31	-
2	3 (16.67%)	12	-	22	2 (11.11%)	32	2 (11.11%)
3	6 (33.33%)	13	13 (72.22%)	23	2 (11.11%)	33	6 (33.33%)
4	5 (27.78%)	14	-	24	6 (33.33%)	34	4 (22.22%)
5	5 (27.78%)	15	2 (11.11%)	25	13 (72.22%)	35	2 (11.11%)
6	1 (5.56%)	16	-	26	-	36	13 (72.22%)
7	1 (5.56%)	17	5 (27.78%)	27	18 (100.00%)	37	17 (94.44%)
8	9 (50.00%)	18	7 (38.89%)	28	-	38	1 (5.56%)
9	12 (66.67%)	19	7 (38.89%)	29	6 (33.33%)	39	18 (100.00%)
10	2 (11.11%)	20	4 (22.22%)	30	17 (94.44%)		

PASSAGE C

1	-	11	-	21	7 (38.89%)	31	8 (44.44%)
2	3 (16.67%)	12	2 (11.11%)	22	3 (16.67%)		
3	5 (27.78%)	13	4 (22.22%)	23	-		
4	10 (55.56%)	14	3 (16.67%)	24	11 (61.11%)		
5	-	15	10 (55.56%)	25	1 (5.56%)		
6	6 (33.33%)	16	1 (5.56%)	26	1 (5.56%)		
7	-	17	3 (16.67%)	27	4 (22.22%)		
8	6 (33.33%)	18	-	28	6 (33.33%)		
9	9 (50.00%)	19	8 (44.44%)	29	-		
10	-	20	3 (16.67%)	30	7 (38.89%)		

(v) COMBINED (WEIGHTED SCORES)

PASSAGE A

1	-	11	1 (1.05%)	21	1 (1.05%)	31	1 (1.05%)
2	1 (1.05%)	12	24 (25.26%)	22	-	32	24 (25.26%)
3	10 (10.53%)	13	-	23	5 (5.26%)	33	-
4	6 (6.32%)	14	23 (24.21%)	24	37 (38.95%)	34	1 (1.05%)
5	6 (6.32%)	15	-	25	8 (8.42%)	35	2 (2.11%)
6	3 (3.16%)	16	31 (32.63%)	26	4 (4.21%)	36	10 (10.53%)
7	9 (9.47%)	17	22 (23.16%)	27	67 (70.53%)		
8	24 (25.26%)	18	35 (36.84%)	28	4 (4.21%)		
9	34 (35.79%)	19	-	29	18 (18.95%)		
10	6 (6.32%)	20	95 (100.00%)	30	3 (3.16%)		

PASSAGE B

1	-	11	71 (78.89%)	21	-	31	-
2	3 (1.20%)	12	-	22	75 (83.33%)	32	2 (2.22%)
3	14 (15.56%)	13	17 (18.89%)	23	2 (2.22%)	33	30 (33.33%)
4	9 (10.00%)	14	-	24	46 (51.11%)	34	44 (48.89%)
5	9 (10.00%)	15	6 (6.67%)	25	29 (32.22%)	35	2 (2.22%)
6	1 (1.11%)	16	-	26	-	36	13 (14.44%)
7	1 (1.11%)	17	55 (61.11%)	27	18 (20.00%)	37	17 (18.89%)
8	21 (23.33%)	18	15 (16.67%)	28	-	38	1 (1.11%)
9	20 (22.22%)	19	16 (17.78%)	29	6 (6.67%)	39	18 (20.00%)
10	2 (2.22%)	20	4 (4.44%)	30	21 (23.33%)		

PASSAGE C

1	-	11	70 (77.78%)	21	10 (11.11%)	31	29 (32.22%)
2	3 (3.33%)	12	2 (2.22%)	22	59 (65.56%)		
3	5 (5.56%)	13	4 (4.44%)	23	-		
4	19 (21.11%)	14	3 (3.33%)	24	15 (16.67%)		
5	-	15	10 (11.11%)	25	1 (1.11%)		
6	9 (10.00%)	16	64 (71.11%)	26	1 (1.11%)		
7	-	17	12 (13.33%)	27	32 (35.56%)		
8	28 (31.11%)	18	5 (5.56%)	28	37 (41.11%)		
9	12 (13.33%)	19	20 (22.22%)	29	-		
10	-	20	3 (3.33%)	30	10 (11.11%)		

APPENDIX 8

EXPERIMENT 3 : STATISTICAL ANALYSES

(a) Spearman Rank Correlation Coefficient r_s (1)

(b) Spearman Rank Correlation Coefficient r_s (2)

(c) Spearman Rank Correlation Coefficient r_s (3)

(d) Mann-Whitney U-Test (1)

(e) Mann-Whitney U-Test (2)

(a) SPEARMAN RANK CORRELATION COEFFICIENT r_s (1)

			SCORES		RANKS			
NO	PASS-AGE	SENT-ENCE	VMI	DIV-ISION	VMI	DIV-ISION	d_i	d_i^2
1	A	20	0.50	1	2	2	0	0
2	B	22	0.63	1	1	2	-1	1
3	C	16	0.09	1	13	2	+11	121
4	A	27	0.27	2	4.5	6	-1.5	2.25
5	B	11	0.27	2	4.5	6	-1.5	2.25
6	B	17	0.16	2	7.5	6	+1.5	2.25
7	C	11	0.40	2	3	6	-3	9
8	C	22	0.08	2	14	6	+8	64
9	A	18	0.16	3	7.5	9	-1.5	2.25
10	A	12	0.16	4	7.5	14	-6.5	42.25
11	A	16	0.14	4	10.5	14	-3.5	12.25
12	A	17	0.07	4	16.5	14	+2.5	6.25
13	B	8	0.14	4	10.5	14	-3.5	12.25
14	B	9	0.07	4	16.5	14	+2.5	6.25
15	B	27	0.07	4	16.5	14	+2.5	6.25
16	B	39	0.12	4	12	14	-2	4
17	C	24	0.16	4	7.5	14	-6.5	42.25
18	A	32	0.07	4	16.5	14	+2.5	6.25

$\Sigma d_i^2 = 342$

(b) SPEARMAN RANK CORRELATION COEFFICIENT r_s (2)

NO	PASS-AGE	SENT-ENCE	SCORES		RANKS		d_i	d_i^2
			VMI	WEIGHTED %	VMI	WEIGHTED %		
1	A	20	0.50	100.00	2	1	-1	1
2	A	27	0.27	70.53	5.5	6	+0.5	0.25
3	A	12	0.16	25.26	10.5	16.5	+6	36
4	A	18	0.16	36.84	10.5	10	-0.5	0.25
5	A	16	0.14	32.63	14	14	0	0
6	A	9	0.09	35.79	17.5	11	-6.5	42.25
7	A	23	0.08	5.26	20	31	+11	121
8	A	24	0.08	38.95	9	24	+15	225
9	A	17	0.07	23.16	25.5	19	-6.5	42.25
10	A	26	0.07	4.21	25.5	33	+7.5	56.25
11	A	32	0.07	25.26	25.5	16.5	-9	81
12	A	36	0.06	10.53	30.5	27	-3.5	12.25
13	A	33	0.05	-	32.5	34	+1.5	2.25
14	B	22	0.63	83.33	1	2	+1	1
15	B	33	0.35	33.33	4	13	+9	81
16	B	11	0.27	78.89	5.5	3	-2.5	6.25
17	B	17	0.16	61.11	10.5	8	-2.5	6.25
18	B	8	0.14	23.33	14	18	+4	16
19	B	39	0.12	20.00	16	21.5	+5.5	30.25
20	B	9	0.07	22.22	25.5	20	-5.5	30.25
21	B	15	0.07	6.67	25.5	29.5	+4	16
22	B	18	0.07	16.67	25.5	23.5	-2	4
23	B	27	0.07	20.00	25.5	21.5	-4	16
24	B	5	0.06	10.00	30.5	28	-2.5	6.25
25	B	29	0.05	6.67	32.5	29.5	-3	9
26	C	11	0.40	77.78	3	4	+1	1
27	C	8	0.24	31.11	7	15	+8	64
28	C	13	0.18	4.44	8	32	+24	576
29	C	24	0.16	16.67	10.5	23.5	+13	169
30	C	21	0.14	11.11	14	25.5	+11.5	132.25
31	C	16	0.09	71.11	17.5	5	-12.5	156.25
32	C	22	0.08	65.56	20	7	-13	169
33	C	27	0.07	35.56	25.5	12	-13.5	182.25
34	C	30	0.04	11.11	34	25.5	-8.5	72.25

$\leq d_i^2 = 2364$

(c) SPEARMAN RANK CORRELATION COEFFICIENT r_s (3)

			SCORES		RANKS			
NO	PASS-AGE	SENT-ENCE	NO OF PROPS	DIVI-SION	NO OF PROPS	DIVI-SION	d_i	d_i^2
1	A	20	43	1	1	2	-1	1
2	B	22	30	1	3	2	+1	1
3	C	16	40	1	2	2	0	0
4	A	9	21	2	6.5	7.5	-1	1
5	A	27	27	2	5	7.5	-2.5	6.25
6	B	11	9	2	15	7.5	+7.5	56.25
7	B	17	10	2	13.5	7.5	+6	36
8	B	24	29	2	4	7.5	-3.5	12.25
9	B	34	10	2	13.5	7.5	+6	36
10	C	11	11	2	11.5	7.5	+4	16
11	C	22	3	2	23.5	7.5	+16	256
12	A	18	3	3	23.5	12.5	+11	121
13	C	28	11	3	11.5	12.5	-1	1
14	A	12	14	4	9	24	-15	225
15	A	14	5	4	19	24	-5	25
16	A	16	3	4	23.5	24	-0.5	0.25
17	A	17	2	4	28	24	+4	16
18	A	29	21	4	6.5	24	-17.5	306.25
19	A	32	13	4	10	24	-14	196
20	B	8	2	4	28	24	+4	16
21	B	9	3	4	23.5	24	-0.5	0.25
22	B	13	7	4	17.5	24	-6.5	42.25
23	B	25	1	4	31	24	+7	49
24	B	27	1	4	31	24	+7	49
25	B	30	7	4	17.5	24	-6.5	42.25
26	B	36	0	4	33.5	24	+9.5	90.25
27	B	37	4	4	20	24	-4	16
28	B	39	1	4	31	24	+7	49
29	C	4	8	4	16	24	-8	64
30	C	9	3	4	23.5	24	-0.5	0.25
31	C	15	2	4	28	24	+4	16
32	C	19	3	4	23.5	24	-0.5	0.25
33	C	24	20	4	8	24	-16	256
34	C	31	0	4	33.5	24	+9.5	90.25

$$\Sigma d_i^2 = 2093$$

(d) MANN - WHITNEY U-TEST (1)

SET 1 : WITHOUT EXTRINSIC SIGNALS

PASS-AGE	SENT-ENCE	%	RANK
B	9	66.67	20
A	17	63.16	18.5
A	29	63.16	18.5
A	12	57.89	16
B	17	55.56	12
B	24	55.56	12
B	34	55.56	12
C	15	55.56	12
C	16	55.56	12
C	22	55.56	12
B	8	50.00	7.5
A	16	47.37	5
A	32	47.37	5
C	19	44.44	2.5
C	31	44.44	2.5

RANK TOTAL (R_1) = 167.5

SET 2 : WITH EXTRINSIC SIGNALS

PASS-AGE	SENT-ENCE	%	RANK
A	20	100.00	31
B	27	100.00	31
B	39	100.00	31
B	30	94.44	28.5
B	37	94.44	28.5
B	11	83.33	27
C	11	77.78	26
B	13	72.22	23.5
B	22	72.22	23.5
B	25	72.22	23.5
B	36	72.22	23.5
A	27	68.42	21
C	24	61.11	17
C	4	55.56	12
C	9	50.00	7.5
A	14	47.37	5
A	18	42.11	1

RANK TOTAL (R_2) = 360.5

(a) MANN-WHITNEY U-TEST (2)

SET 1 : WITHOUT INTRINSIC SIGNALS

PASS-AGE	SENT-ENCE	%	RANK
B	27	100.00	31
B	39	100.00	31
B	13	72.22	23.5
A	27	68.42	21
C	24	61.11	17
B	34	55.56	12
C	4	55.56	12
C	15	55.56	12
C	9	50.00	7.5
A	32	47.37	5
A	18	42.11	1

RANK TOTAL (R_1) = 173

SET 2 : WITH INTRINSIC SIGNALS

PASS-AGE	SENT-ENCE	%	RANK
B	20	100.00	31
B	30	94.44	28.5
B	37	94.44	28.5
B	11	83.33	27
C	11	77.78	26
B	22	72.22	23.5
B	25	72.22	23.5
B	36	72.22	23.5
B	9	66.67	20
A	17	63.16	18.5
A	29	63.16	18.5
A	12	57.89	16
B	17	55.56	12
B	24	55.56	12
C	16	55.56	12
C	22	55.56	12
B	8	50.00	7.5
A	14	47.37	5
A	16	47.37	5
C	19	44.44	2.5
C	31	44.44	2.5

RANK TOTAL (R_2) = 355

APPENDIX 9

TEXTBOOKS USED FOR THE ANALYSIS OF THEME SENTENCE LENGTHS

TEXTBOOKS USED FOR THE ANALYSIS OF THEME SENTENCE LENGTHS

1. Hooson, D. J. M. (1966) The Soviet Union University of London Press.

2. Houston, J. M. (1964) The Western Mediterranean World Longman

3. Mellor, R. E. H. (1965) The Geography of the USSR MacMillan.

4. Morgan, W. T. W. (1972) East Africa Longman.

5. Paterson, J. H. (1965) North America : A Regional Geography Oxford University Press.

6. Shorter, A. H., W. L. D. Ravenhill & K. J. Gregory (1969) Southwest England Nelson.

7. Sinnhuber, K. A. (1961) Germany : Its Geography and Growth Murray.

8. Tregear, T. R. (1965) A Geography of China University of London Press.

9. Walker, F. (1972) The Bristol Region Nelson.

10. Watson, J. W. (1963) North America : Its Countries and Regions Longman.

APPENDIX 10

EXPERIMENT 4 : THE EXPERIMENTAL PASSAGES

(a) Wide Format, Version A (4 paragraphs)
(b) Wide Format, Version B (8 paragraphs)
(c) Wide Format, Version C (12 paragraphs)
(d) Normal Format, Version A (4 paragraphs)
(e) Normal Format, Version B (8 paragraphs)
(f) Normal Format, Version C (12 paragraphs)
(g) Columns Format, Version A (4 paragraphs)
(h) Columns Format, Version B (8 paragraphs)
(i) Columns Format, Version C (12 paragraphs)

EXPERIMENT 4

(a) Wide Format, Version A

Every geologist is familiar with the erosion cycle. No sooner has an area of land been raised above sea level than it is subject to the erosive forces of nature. The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets, and then into rivers and out to sea. The frost freezes the rain water in cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust. Blocks of rock dislodged at high levels are brought down by the force of gravity. Alternate heating and cooling of bare rock surfaces causes their disintegration. In the arid regions of the world the wind is a powerful force in removing material from one area to another.

All this is natural. But Nature has also provided certain defensive forces. Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks. Slowly but surely, different types of soil with differing profiles evolve, the main types depending primarily on the climate. The protective soil covering, once it is formed, is held together by the vegetation. Grass and herbaceous plants, with long, branching, tenuous roots, hold firmly together the surface particles. The same is true with the forest cover. The heaviest tropical downpours beating on the leaves of the giant trees reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil. The soil, thus protected by grass, herbs or trees, furnishes a quiet habitat for a myriad varied organisms : earthworms that importantly modify the soil, bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants. Chemical action is constantly taking place : soil acids attack mineral particles, and salts in solution move from one layer in the soil to another. We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions and the parent rock substances gradually develops. It is a long process, and in many areas there has not been sufficient time, in the geological sense, for completion of the process : the soils are 'immature'.

Now let us consider what happens when man, the pioneer, comes along. He ploughs up the natural grasslands. He removes the numerous branching roots that have held together the surface particles, so that they are now easily moved by the action of rain and wind. In ploughing, he has mixed together the surface layers and provided his crops with a medium in which they can grow and develop, and in which natural plant food is present in varying degree, or which may be supplemented by animal or chemical manures. But he has also exposed the soil to the action of the atmosphere. In some parts of the world exposure is useful : in a cold climate the breaking up of the clods of clay by frost action is beneficial. At the other extreme, as we have already pointed out, in tropical climates, by exposure to the atmosphere such rapid chemical action may have been set up that the natural plant food in the soil is quickly destroyed.

Broadly speaking,the position is worse when the natural vegetation is forest. Not only does man, the pioneer, rob the soil of its source of humus, but he exposes a soil quite unused to the direct rays of the sun and the direct fall of the rain to the immediate influence of both. For example, much of the upland and west coast of Scotland was once forested with the beautiful Scots fir. The heavy rain, falling on the close pine woods, trickled gradually to the ground and soaked into the soil. Much was evaporated from the leaves, and the floor of the forest, covered with pine leaves, remained comparatively dry, supporting a sparce cover or undergrowth, of various shade-loving plants or low shrubs, such as bilberry or heather. When the forests were cut down the heavy rain fell straight on to the surface soil more rapidly than it could drain away. Especially where there was no steep slope the water was held up, and moisture-loving plants began to flourish, particularly spaghnum, or bog moss. Once the spaghnum was established it acted as a sponge. True, it prevented soil erosion, but it grew and grew until great thicknesses of moss blanketed the whole countryside. Thus, huge stretches of bog land, that known to the botanists as 'blanket bog', extending over wide areas of Scotland and Ireland, are directly due to man's action.

Every geologist is familiar with the erosion cycle. No sooner has an area of land been raised above sea level than it becomes subject to the erosive forces of nature.The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets and then into rivers and out to sea. The frost freezes the rain water in cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust. Blocks of rock dislodged at high levels are brought down by the force of gravity. Alternate heating and cooling of bare rock surfaces causes their disintegration. In the arid regions of the world the wind is a powerful force in removing material from one area to another.

All this is natural. But Nature has also provided certain defensive forces. Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks. Slowly but surely, different types of soil, with differing profiles, evolve, the main types depending primarily on the climate.

The protective soil covering, once it is formed, is held together by the vegetation. Grass and herbaceous plants, with long, branching, tenuous roots, hold firmly together the surface particles. The same is true with the forest cover. The heaviest tropical downpours, beating on the leavesof the giant trees, reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil. The soil, thus protected by grass, herbs or trees, furnishes a quiet habitat for a myriad varied organisms : earthworms, that importantly modify the soil, bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants. Chemical action is constantly taking place : soil acids attack mineral particles and salts in solution move from one layer in the soil to another.

We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions and the parent rock substances develops. It is a long process, and in many areas there has not been sufficient time, in the geological sense, for completion of the process : the soils are 'immature'.

Now let us consider what happens when man, the pioneer, comes along. He ploughs up the natural grasslands. He removes the various branching roots that have held together the surface particles, so that they are now easily moved by the action of rain and wind.

In ploughing he has mixed together the surface layers, and provided his crops with a medium in which they can grow and develop, and in which natural plant food is present in varying degree, or which can be supplemented by animal or chemical manures. But he has also exposed the soil to the action of the atmosphere. In some parts of the world exposure is useful : in a cold climate the breaking up of the clods of clay by frost action is beneficial At the other extreme, as we have already pointed out, in tropical climates, by exposure to th atmosphere such rapid chemical action is set up that the natural plant food in the soil is quickly destroyed.

Broadly speaking, the position is worse when the natural vegetation is forest. Not only does man, the pioneer, rob the soil of its source of humus, the fallen leaves, but he exposes a soil quite unused to the direct rays of the sun and the direct fall of the rain to the immediate influence of both.

For example, much of the upland and west coast of Scotland was once forested with the beautiful Scots fir. The heavy rain, falling on the close pine woods, trickled gradually to the ground and soaked into the soil. Much was evaporated from the leaves, and the floor of the forest, covered with pine leaves, remained comparatively dry, supporting a sparce cover or undergrowth of various shade-loving plants or low shrubs such as bilberry and heather. When the forests were cut down the heavy rain fell straight on to the surface soil more rapidly than it could drain away. Especially where there was no steep slope the water was held up, and moisture-loving plants began to flourish, particularly spaghnum, or bog moss. Once the spaghnum was established it acted as a sponge. True, it prevented soil erosion, but it grew and grew until great thicknesses of moss blanketed the whole countryside. Thus, huge stretches of bog land, that known to the botanists as 'blanket bog', extending over wide areas of Scotland and Ireland, are directly due to man's action.

Every geologist is familiar with the erosion cycle. No sooner has an area of land been raised above sea level than it become subject to the erosive forces of nature. The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets, and then into rivers and out to sea. The frost freezes the rain water in cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust. Blocks of rock dislodged at high levels are brought down by the force of gravity. Alternate heating and cooling of bare rock surfaces causes their disintegration. In the arid regions of the world the wind is a powerful force in removing material from one area to another.

All this is natural. But Nature has also provided ceratin defensive forces. Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks. Slowly but surely, different types of soil, with differing profiles, evolve, the main types depending primarily on the climate.

The protective soil covering, once it is formed, is held together by the vegetation. Grass and herbaceous plants, with long, branching, tenuous roots, hold firmly together the surface particles. The same is true with the forest cover. The heaviest tropical downpours, beating on the leavesof the giant trees, reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil.

The soil, thus protected by grass, herbs and trees, furnishes a quiet habitat for a myriad varied organisms : earthworms that importantly modify the soil, bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants.

Chemical action is constantly taking place : soil acids attack mineral particles and salts in solution move from one layer in the soil to another.

We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions and the parent rock substances develops. It is a long process, and in many areas there has not been sufficient time, in the geological sense, for completion of the process : the soils are 'immature'.

Now let us consider what happens when man, the pioneer, comes along. He ploughs up the natural grasslands. He removes the various branching roots that have held together the surface particles, so that they are now easily moved by the action of rain and wind.

In ploughing he has mixed together the surface layers, and provided his crops with a medium in which they can grow and develop, and in which natural plant food is present in varying degrees, or which can be supplemented by animal or chemical manures. But he has also exposed the soil to the action of the atmosphere.

In some parts of the world exposure is useful : in a cold climate the breaking up of the clods of clay by frost action is beneficial. At the other extreme, as we have already pointed out, in tropical climates, by exposure to the atmosphere such rapid chemical action is set up that the natural plant food in the soil is quickly destroyed.

Broadly speaking, the position is worse when the natural vegetationis forest. Not only does man, the pioneer, rob the soil of its source of humus, the fallen leaves, but he exposes a soil quite unused to the direct rays of the sun and the direct fall of the rain to the immediate influence of both.

For example, much of the upland and west coast of Scotland was once forested with the beautiful Scots fir. The heavy rain, falling on the close pine woods, trickled gradually to the ground and soaked into the soil. Much was evaporated from the leaves, and the floor of the forest, covered with pine leaves, remained comparatively dry, supporting a sparse cover or undergrowth of various shade-loving plants, or low shrubs, such as bilberry and heather.

When the forests were cut down the heavy rain fell straight on to the surface soil more rapidly than it could drain away. Especially where there was no steep slope the water was held up and moisture-loving plants began to flourish, particularly spaghnum, or bog moss. Once the spaghnum was established it acted as a sponge. True, it prevented soil erosion, but it grew and grew until great thicknesses of moss blanketed the whole countryside. Thus, huge stretches of bog land, that known to the botanists as 'blanket bog', extending over wide areas of Scotland and Ireland, are directly due to man's action.

Every geologist is familiar with the erosion cycle. No sooner has an area of land been raised above sea level than it is subject to the erosive forces of nature. The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets and then into rivers and out to sea. The frost freezes the rain water into cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust. Blocks of rock dislodged at high levels are brought down by the force of gravity. Alternate heating and cooling of bare rock surfaces causes their disintegration. In the arid regions of the world the wind is a powerful force in removing material from one area to another.

All this is natural. But Nature has also provided certain defensive forces. Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks. Slowly but surely different types of soil with differing profiles evolve, the main types depending primarily on the climate. The protective soil covering, once it is formed, is held together by the vegetation. Grass and herbaceous plants, with long, branching, tenuous roots, hold firmly together the surface particles. The same is true with the forest cover. The heaviest tropical downpours beating on the leaves of the giant trees reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil. The soil, thus protected by grass, herbs or trees; furnishes a quiet habitat for a myriad varied organisms : earthworms that importantly modify the soil, bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants. Chemical action is constantly taking place : soil acids attack mineral particles, and salts in solution move from one layer in the soil to another. We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions and the parent rock substances gradually develops. It is a long process, and in many areas there has not been sufficient time, in the geological sense, for completion of the process : the soils are 'immature'.

Now let us consider what happens when man, the pioneer, comes along. He ploughs up the natural grasslands. He removes the numerous branching roots that have held together the surface particles, so that they are now easily moved by the action of rain and wind. In ploughing, he has mixed together the surface layers and provided his crops with a medium in which they can grow and develop, and in which natural plant food is present in varying degree, or which may be supplemented by animal or chemical manures. But he has also exposed the soil to the action of the atmosphere. In some parts of the world exposure is useful : in a cold climate the breaking up of the clods of clay by frost action is beneficial. At the other extreme, as we have already pointed out, in tropical climates, by exposure to the atmosphere such rapid chemical action may have been set up that the natural plant food in the soil is quickly destroyed.

Broadly speaking, the position is worse when the natural vegetation is forest. Not only does man, the pioneer, rob the soil of its source of humus, but he exposes a soil quite unused to the direct rays of the sun and the direct fall of the rain to the immediate influence of both. For example, much of the upland and west coast of Scotland was once forested with the beautiful Scots fir. The heavy rain, falling on the close pine woods, trickled gradually to the ground and soaked into the soil. Much was evaporated from the leaves, and the floor of the forest, covered with pine leaves, remained comparatively dry, supporting a sparce cover of undergrowth of various shade-loving plants or low shrubs, such as bilberry or heather. When the forests were cut down the heavy rain fell straight on to the surface soil more rapidly than it could drain away. Especially where there was no steep slope the water was held up, and the moisture-loving plants began to flourish, particulary spaghnum, or bog moss. Once the spaghnum was established it acted as a sponge. True, it prevented soil erosion, but it grew and grew until great thicknesses of moss blanketed the whole countryside. Thus, huge stretches of bog land, that known to the botanists as 'blanket bog', extending over wide areas of Scotland and Ireland, are directly due to man's action.

EXPERIMENT 4 (a) Normal Format, Version 8

Every geologist is familiar with the erosion cycle. No sooer has an area of land been raised above sea level than it becomes subject to the erosive forces of nature. The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets and then into rivers and out to sea. The frost freezes the rain water in cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust. Blocks of rock dislodged at high levels are brought down by the force of gravity. Alternate heating and cooling of bare rock surfaces causes their disintegration. In the arid regions of the world the wind is a powerful force in removing material from one area to another.

All this is natural. But Nature has also provided certain defensive forces. Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks. Slowly but surely, different types of soil, with differing profiles, evolve, the main types depending primarily on the climate.

The protective soil covering, once it is formed, is held together by the vegetation. Grass and herbaceous plants, with long, branching, tenuous roots, hold firmly together the surface particles. The same is true with the forest cover. The heaviest tropical downpours, beating on the leaves of the giant trees, reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil. The soil, thus protected by grass, herbs or trees, furnishes a quiet habitat for a myriad varied organisms : earthworms, that importantly modify the soil, bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants. Chemical action is constantly taking place : soil acids attack mineral particles and salts in solution move from one layer in the soil to another.

We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions and the parent rock substances develops. It is a long process, andin many areas there has not been sufficient time, in the geological sense, for completion of the process : the soils are 'immature'.

Now let us consider what happens when man, the pioneer, comes along. He ploughs up the natural grasslands. He removes the various branching roots that have held together the surface particles, so that they are now easily moved by the action of rain and wind.

In ploughing he has mixed together the surface layers, and provided his crops with a medium in which they can grow and develop, and in which natural plant food is present in varying degree, or which can be supplemented by animal or chemical manures. But he has also exposed the soil to the action of the atmosphere. In some partsof the world exposure is useful : in a cold climate the breaking up of the clods of clay by frost action is beneficial. At the other extreme, as we have already pointed out, in tropical climates, by exposure to the atmosphere such rapid chemical action is set up that the natural plant food in the soil is quickly destroyed.

Broadly speaking, the position is worse when the natural vegetation is forest. Not only does man, the pioneer, rob the soil of its source of humus, the fallen leaves, but he exposes a soil quite unused to the direct rays of the sun and the direct fall of the rain to the immediate influence of both.

For example, much of the upland and west coast of Scotland was once forested with the beautiful Scots fir. The heavy rain, falling on the close pine woods, trickled gradually to the ground and soaked into the soil. Much was evaporated from the leaves, and the floor of the forest, covered with pine leaves, remained comparatively dry, supporting a sparse cover of undergrowth of various shade-loving plants or low shrubs, such as bilberry and heather. When the forests were cut down the rain fell straight on to the surface soil more rapidly than it could drain away. Especially where there was no steep slope the water was held up, and moisture-loving plants began to flourish, particularly spaghnum, or bog moss. Once the spaghnum was established it acted as a sponge. True, it prevented soil erosion, but it grew and grew until great thicknesses of moss blanketed the whole countryside. Thus, huge stretches of bog land, that known to the botanists as 'blanket bog', extending over wide areasof Scotland and Ireland, are directly due to man's action.

EXPERIMENT 4 (f) Normal Format, Version C

Every geologist is familiar with the erosion cycle. No sooner has an area of land been raised above sea level than it becomes subject to the erosive forces of nature. The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets, and then into rivers and out to sea. The frost freezes the rain water in cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust. Blocks of rock dislodged at high levels are brought down by the force of gravity. Alternate heating and cooling of bare rock surfaces causes their disintegration. In the arid regions of the world the wind is a powerful force in removing material from one area to another.

All this is natural. But Nature has also provided certain defensive forces. Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks. Slowly but surely, different types of soil, with differing profiles, evolve, the main types depending primarily on the climate.

The protective soil covering, once it is formed, is held together by the vegetation. Grass and herbaceous plants, with long, branching, tenuous roots, hold firmly together the surface particles. The same is true with the forest cover. The heaviest tropical downpours, beating on the leaves of the giant trees, reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil.

The soil, thus protected by grass, herbs or trees, furnishes a quiet habitat for a myriad varied organisms : earthworms that importantly modify the soil, bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants.

Chemical action is constantly taking place : soil acids attack mineral particles, and salts in solution move from one layer in the soil to another.

We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions and the parent rock substances develops. It is a long process, and in many areas there has not been sufficient time, in the geological sense, for completion of the process : the soils are 'immature'.

Now let us consider what happens when man, the pioneer, comes along. He ploughs up the natural grasslands. He removes the various branching roots that have held together the surface particles, so that they are now easily moved by the action of rain and wind.

In ploughing, he has mixed together the surface layers, and provided his crops with a medium in which they can grow and develop, and in which natural plant food is present in varying degrees, or which can be supplemented by animal or chemical manures. But he has also exposed the soil to the action of the atmosphere.

In some parts of the world exposure is useful : in a cold climate the breaking up of the clods of clay by frost action is beneficial. At the other extreme, as we have already pointed out, in tropical climates, by exposure to the atmosphere such rapid chemical action is set up that the natural plant food in the soil is quickly destroyed.

Broadly speaking, the position is worse when the natural vegetation is forest. Not only does man, the pioneer, rob the soil of its source of humus, the fallen leaves, but he exposes a soil quite unused to the direct rays of the sun and the direct fall of the rain to the immediate influence of both.

For example, much of the upland and west coast of Scotland was once forested with the beautiful Scots fir. The heavy rain, falling on the close pine woods, trickled gradually to the ground and soaked into the soil. Much was evaporated from the leaves, and the floor of the forest, covered with pine leaves, remained comparatively dry, supporting a sparse cover or undergrowth of various shade-loving plants, or low shrubs, such as bilberry and heather.

When the forests were cut down the heavy rain fell straight on to the surface soil more rapidly than it could drain away. Especially where there was no steep slope the water was held up, and moisture-loving plants began to flourish, particularly spaghnum, or bog moss. Once the spaghnum was established it acted as a sponge. True, it prevented soil erosion, but it grew and grew until great thickness of moss blanketed the whole countryside. Thus, huge stretches of bogland, that known to the botanists as 'blanket bog', extending over wide areas of Scotland and Ireland, are directly due to man's action.

(g) Columns Format, Version A

Every geologist is familiar with the erosion cycle. No sooner has an area of land been raised above sea-level than it becomes subject to the erosive forces of nature. The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets and then into rivers and out to sea. The frost freezes the rain water in cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust. Blocks of rock dislodged at high levels are brought down by the force of gravity. Alternate heating and cooling of bare rock surfaces causes their disintegration. In the arid regions of the world the wind is a powerful force in removing material from one area to another.

All this is natural. But Nature has also provided certain defensive forces. Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks. Slowly but surely different types of soil with differing profiles evolve, the main types depending on the climate. The protective soil covering, once it is formed, is held together by the vegetation. Grass and herbaceous plants, with long, branching tenuous roots, hold firmly together the surface particles. The same is true with the forest cover. The heaviest tropical downpours beating on the leaves of the giant trees reach the ground only as spray, gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil. The soil, thus protected by grass, herbs or trees, furnishes a quiet habitat for a myriad varied organisms : earthworms that importantly modify the soil, bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants. Chemical action is constantly taking place : soil acids attack mineral particles and salts in solution move from one layer in the soil to another. We may sum up by saying that under the natural cover the soil profile proceeds to the climatic conditions and the parent rock substances gradually develops. It is a long process, and in many areas there has not been sufficient time, in the geological sense, for completion of the process: the soils are 'immature'.

Now let us consider what happens when man the pioneer comes along. He ploughs up the natural grasslands. He removes the numerous branching roots that have held together the surface particles, so that they are now easily moved by the action of rain and wind. In ploughing he has mixed together the surface layers and provided his crops with a medium in which they can grow and develop, and in which natural plant food is present in varying degree, or which may be supplemented by animal or chemical manures. But he has also exposed the soil to the action of the atmosphere. In some parts of the world exposure is useful : in a cold climate the breaking up of the clods of clay by frost action is beneficial. At the other extreme, as we have already pointed out, in tropical climates, by exposure to the atmosphere such rapid chemical action may have been set up that the natural plant food in the soil is quickly destroyed.

Broadly speaking, the position is worse when the natural vegetation is forest. Not only does man, the pioneer rob the soil of its humus, the fallen leaves, but he exposes a soil quite unused to the direct rays of the sun and the direct fall of the rain to the immediate influence of both. For example, much of the upland and west coast of Scotland was once forested with the beautiful Scots fir. The heavy rain, falling on the close pine woods, trickled gradually to the ground and soaked into the soil. Much was evaporated from the leaves, and the floor of the forest, covered with pine leaves, remained comparatively dry, supporting a sparse cover or undergrowth of various shade-loving plants or low shrubs, such as bilberry or heather. When the forests were cut down the heavy rain fell straight on to the surface soil more rapidly than it could drain away. Especially where there was no steep slope the water was held up, and moisture-loving plants began to flourish, particularly sphagnum, or bog moss. Once the sphagnum was established it acted as a sponge. True, it prevented soil erosion, but it grew and grew until great thicknesses of moss blanketed the whole countryside. Thus, huge stretches of bogland, that known to the botanists as 'blanket bog', extending over wide areas of Scotland and Ireland, are directly due to man's action.

Every geologist is familiar with the erosion cycle. No sooner has an area of land been raised above sea-level than it becomes subject to the erosive forces of nature. The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets and then into rivers and out to sea. The frost freezes the rain water in cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust. Blocks of rock dislodged at high levels are brought down by the force of gravity. Alternate heating and cooling of bare rock surfaces causes their disintegration. In the arid regions of the world the wind is a powerful force in removing material from one area to another.

All this is natural. But Nature has also provided certain defensive forces. Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks. Slowly but surely different types of soil, with differing profiles, evolve, the main types depending primarily on the climate.

The protective soil covering, once it is formed, is held together by the vegetation. Grass and herbaceous plants, with long, branching tenuous roots, hold firmly together the surface particles. The same is true with the forest cover. The heaviest tropical downpours, beating on the leaves of the giant trees, reach the ground only as spray gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil. The soil, thus protected by grass, herbs or trees, furnishes a quiet habitat for a myriad varied organisms : earthworms that importantly modify the soil, bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants. Chemical action is constantly taking place : soil acids attack mineral particles and salts in solution move from one layer in the soil to another.

We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions and the parent rock substances develops. It is a long process, and in many areas there has not been sufficient time, in the geological sense, for completion of the process : the soils are 'immature'.

Now let us consider what happens when man the pioneer comes along. He ploughs up the natural grasslands. He removes the various branching roots that have held together the surface particles, so that they are now easily moved by the action of rain and wind.

In ploughing, he has mixed together the surface layers, and provided his crops with a medium in which they can grow and develop, and in which natural plant food is present in varying degree, or which can be supplemented by animal or chemical manures. But he has also exposed the soil to the action of the atmosphere. In some parts of the world exposure is useful : in a cold climate the breaking up of the clods of clay by frost action is beneficial. At the other extreme, as we have already pointed out, in tropical climates, by exposure to the atmosphere such rapid chemical action is set up that the natural plant food in the soil is quickly destroyed.

Broadly speaking, the position is worse when the natural vegetation is forest. Not only does man, the pioneer, rob the soil of its source of humus, the fallen leaves, but he exposes a soil quite unused to the direct rays of the sun and the direct fall of the rain to the immediate influence of both.

For example, much of the upland and west coast of Scotland was once forested with the beautiful Scots fir. The heavy rain, falling on the close pine woods, trickled gradually to the ground and soaked into the soil. Much was evaporated from the leaves and the floor of the forest, covered with pine leaves, remained comparatively dry, supporting a sparse cover or undergrowth of various shade-loving plants or low shrubs, such as bilberry or heather. When the forests were cut down the heavy rain fell straight on to the surface soil more rapidly than it could drain away. Especially where there was no steep slope the water was held up, and moisture-loving plants began to flourish, particularly sphagnum, or bog moss. Once the sphagnum was established it acted as a sponge. True, it prevented soil erosion, but it grew and grew until great thicknesses of moss blanketed the whole countryside. Thus, huge stretches of bogland, that known to the botanists as 'blanket bog', extending over wide areas of Scotland and Ireland, are directly due to man's action.

EXPERIMENT 4 (i) Columns Format, Version C

Every geologist is familiar with the erosion cycle. No sooner has an area of land been raised above sea-level than it becomes subject to the erosive forces of nature. The rain beats down on the ground and washes away the finer particles, sweeping them into rivulets and then into rivers and out to sea. The frost freezes the rain water in cracks in the rocks and breaks up even the hardest of the constituents of the earth's crust. Blocks of rock dislodged at high levels are brought down by the force of gravity. Alternate heating and cooling of bare rock surfaces causes their disintegration. In the arid regions of the world the wind is a powerful force in removing material from one area to another.

All this is natural. But Nature has also provided certain defensive forces. Bare rock surfaces are in due course protected by soil, itself dependent initially on the weathering of the rocks. Slowly but surely different types of soil with differing profiles evolve, the main types depending primarily on the climate.

The protective soil covering, once it is formed, is held together by the vegetation. Grass and herbaceous plants, with long, branching tenuous roots, hold firmly together the surface particles. The same is true with the forest cover. The heaviest tropical downpours, beating on the leaves of the giant trees, reach the ground only as spray gently watering the surface layers and penetrating along the long passages provided by the roots to the lower levels of the soil.

The soil, thus protected by grass, herbs or trees, furnishes a quiet habitat for a myriad varied organisms : earthworms that importantly modify the soil, bacteria, active in their work of converting fallen leaves and decaying vegetation into humus and food for the growing plants.

Chemical action is constantly taking place : soil acids attack mineral particles and salts in solution move from one layer in the soil to another.

We may sum up by saying that under the natural cover the soil profile proper to the climatic conditions and the parent rock substances develops. It is a long process, and in many areas there has not been sufficient time in the geological sense for completion of the process : the soils are 'immature'.

Now let us consider what happens when man the pioneer comes along. He ploughs up the natural grasslands. He removes the various branching roots that have held together the surface particles, so that they are now easily moved by the action of rain and wind.

In ploughing, he has mixed together the surface layers, and provided his crops with a medium in which they can grow and develop, and in which natural plant food is present in varying degree, or which can be supplemented by animal or chemical manures. But he has also exposed the soil to the action of the atmosphere.

In some parts of the world exposure is useful : in a cold climate the breaking up of the clods of clay by frost action is beneficial. At the other extreme, as we have already pointed out, in tropical climates, by exposure to the atmosphere such rapid chemical action is set up that the natural plant food in the soil is quickly destroyed.

Broadly speaking, the position is worse when the natural vegetation is forest. Not only does man, the pioneer, rob the soil of its source of humus, the fallen leaves, but he exposes a soil quite unused to the direct rays of the sun and the direct fall of the rain to the immediate influence of both.

For example, much of the upland and west coast of Scotland was once forested with the beautiful Scots fir. The heavy rain, falling on the close pine woods, trickled gradually to the ground and soaked into the soil. Much was evaporated from the leaves, and the floor of the forest, covered with pine leaves, remained comparatively dry, supporting a sparse cover or undergrowth of various shade-loving plants or low shrubs, such as bilberry or heather.

When the forests were cut down the heavy rain fell straight on to the surface soil more rapidly than it could drain away. Especially where there was no steep slope the water was held up, and moisture-loving plants began to flourish, particularly sphagnum, or bog moss. True, it prevented soil erosion, but it grew and grew until great thicknesses of moss blanketed the whole countryside. Thus, huge stretches of bogland, that known to the botanists as 'blanket bog', extending over wide areas of Scotland and Ireland, are directly due to man's action.

APPENDIX 11

EXPERIMENT 4 : STATISTICAL ANALYSES

(a) Analysis of Residuals : Formats

(b) Analysis of Residuals : Versions

(c) Goodman and Kruskal's Lamda Measure of Association

(a) ANALYSIS OF RESIDUALS : FORMATS

(i) WIDE

Standardised Residuals

	1st	2nd	3rd
A	+3.05	-1.33	-1.72
B	-0.53	+3.05	-2.52
C	-2.52	-1.72	+4.25

Variance of Standardised Residuals

	1st	2nd	3rd
A	0.45	0.45	0.45
B	0.45	0.45	0.45
C	0.45	0.45	0.45

(ii) NORMAL

Standardised Residuals

	1st	2nd	3rd
A	+1.63	-1.63	0
B	+0.41	+2.04	-2.45
C	-2.04	-0.41	+2.45

Variance of Standardised Residuals

	1st	2nd	3rd
A	0.45	0.45	0.45
B	0.45	0.45	0.45
C	0.45	0.45	0.45

(iii) COLUMNS

Standardised Residuals

	1st	2nd	3rd
A	-0.41	-1.63	+2.04
B	+1.22	0	-1.22
C	+0.82	+1.63	-0.82

Variance of Standardised Residuals

	1st	2nd	3rd
A	0.45	0.45	0.45
B	0.45	0.45	0.45
C	0.45	0.45	0.45

(b) ANALYSIS OF RESIDUALS : VERSIONS

(i) 'A' VERSION

Standardised Residuals

	1st	2nd	3rd
A	+1.26	+0.16	-1.46
B	+0.37	-0.19	-0.19
C	-1.78	-0.09	+1.92

Variance of Standardised Residuals

	1st	2nd	3rd
A	0.31	0.31	0.31
B	0.57	0.58	0.58
C	0.43	0.44	0.44

(ii) 'B' VERSION

Standardised Residuals

	1st	2nd	3rd
A	-0.84	+0.05	+0.81
B	+1.01	+0.27	-1.30
C	-1.02	-0.99	+2.04

Variance of Standardised Residuals

	1st	2nd	3rd
A	0.41	0.41	0.41
B	0.29	0.29	0.29
C	0.63	0.64	0.64

(iii) 'C' VERSION

Standardised Residuals

	1st	2nd	3rd
A	-1.32	-0.50	+1.84
B	-1.60	-0.24	+1.88
C	+1.66	+0.36	-2.07

Variance of Standardised Residuals

	1st	2nd	3rd
A	0.60	0.61	0.61
B	0.45	0.46	0.46
C	0.26	0.27	0.27

(c) GOODMAN AND KRUSKAL'S LAMDA MEASURE OF ASSOCIATION

FORMULA

$$\lambda = \frac{\sum^{\text{ROWS}} \max(n_{ij}) + \sum^{\text{COLUMNS}} \max(n_{ij}) - \max(n_{.j}) - \max(n_{i.})}{2N - \max(n_{.j}) - \max(n_{i.})}$$

0 = no predictability

1 = complete predictability